

# Do partially disabled people respond to financial incentives to work?\*

Tunga Kantarci<sup>†</sup>, Wietse Mesman<sup>‡</sup> and Jan-Maarten van Sonsbeek<sup>§</sup>

3 September 2023

Preliminary version

## Abstract

In the Netherlands partially disabled individuals face financial incentives to increase labor participation after 3 to up to 38 months from the first day of claiming disability insurance (DI). The incentive induces a negative income effect on labor supply. By the design of the DI scheme, the income effect is larger for those who earn higher wages before falling sick. Using administrative data on all partially disabled individuals and taking a local randomization approach to regression discontinuity, we analyze the labor participation effect of the financial incentives across pre-sickness wage groups. Partially disabled individuals who earn at most the minimum wage before falling sick increase labor participation by 3.7 percentage points, while those who earn at least 2.5 times of the minimum wage increase labor participation by 7.4 percentage points. The effects are persistent in the long run, substantially heterogeneous across labor market and disability characteristics, and robust to alternative identification assumptions.

## 1 Introduction

In 2006 the Netherlands reformed its DI scheme which led to a strong and persistent decrease in annual inflow into DI (Van Sonsbeek and Gradus, 2013; Kantarci et al., 2023). However, the inflow by partially disabled individuals increased from 0.23% of the working population in 2006 to 0.53% in 2022 (see Figure 1). There are several reasons for this increase, of which the

---

\*This research is supported by Netspar under grant number LMVP 2014.03. Its contents are the sole responsibility of the authors. We thank the Employee Insurance Agency (UWV), and in particular Lucien Rondagh, Willy van den Berk, Carla van Deursen, and Roel Ydema, for providing the disability data. We thank Hans Bloemen, Tobias Klein, Ziwei Rao, Arthur van Soest, and participants of the Netspar Pension Day 2021 and the Netspar International Pension Workshop 2023 for their helpful comments and suggestions. Results are based on calculations by the authors using non-public microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information: [microdata@cbs.nl](mailto:microdata@cbs.nl).

<sup>†</sup>Department of Econometrics and Operations Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, and Netspar (e-mail: [kantarci@tilburguniversity.edu](mailto:kantarci@tilburguniversity.edu))

<sup>‡</sup>Department of Economics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, and Netspar (e-mail: [w.r.b.mesman@tilburguniversity.edu](mailto:w.r.b.mesman@tilburguniversity.edu))

<sup>§</sup>Department of Labor and Knowledge, Netherlands Bureau for Economic Policy Analysis, P.O. Box 80510, 2508 GM The Hague, The Netherlands, and Netspar (e-mail: [j.m.van.sonsbeek@cpb.nl](mailto:j.m.van.sonsbeek@cpb.nl))

most important are the increasing number of older workers and workers without a permanent contract, who both have a higher DI risk in the insured population. During the period from 2006 to 2019, the Netherlands spent 3.01% of its GDP on incapacity, well above the OECD average of 2.01% during the same period ([OECD, 2023](#)).

As the number of DI recipients and public spending on DI grow in many Western countries, reducing the size of DI schemes is important to policy makers. The size of a DI scheme can be reduced by either reducing inflow into the scheme or increasing outflow from it. Inflow can be reduced by stricter screening during the sickness insurance period, or stricter eligibility criteria for DI at the end of it, so that fewer insured workers qualify for DI. DI benefits can also be made less generous so that fewer workers seek them. Outflow can be increased by stronger reintegration measures ([Autor et al., 2016](#)). Financial incentives can both reduce inflow and increase outflow.

A substantial body of literature has developed on the causal impact of financial incentives on reducing inflow and increasing outflow. Studies that focus on DI inflow investigate the impact of changing the benefit level. [Bound and Burkhauser \(1999\)](#) review the studies from the 1970s to 1990s in the US and consistently find that increasing (decreasing) the benefit level increases (decreases) benefit claims and awards, and decreases (increases) labor participation, although sizes of the estimated effects vary considerably. [Gruber \(2000\)](#), [Campolieti \(2004\)](#), and [Favre et al. \(2021\)](#) investigate the impact of the changes in the generosity of DI benefits in the Canadian DI scheme, and [Mullen and Staubli \(2016\)](#) analyze this in Austria. In the 1980s, [Gruber](#) finds that increasing the benefit level leads to a large decrease in labor force participation among older workers. In line with this, [Favre et al.](#) find a large increase in DI claims for the same benefit increase. However, in the 1970s, [Campolieti](#) finds that increasing the benefit level does not lead to a significant increase in DI claims or decrease in labor force participation among older workers. [Mullen and Staubli](#) also find strong responses of DI applications to decreases and increases in benefit levels, but the more recent responses, after the system has become stricter over time, seem to be smaller. This suggests that making the DI scheme stricter has weakened the impact of benefit generosity on benefit use.

To incentivize outflow from DI, many DI schemes allow earnings up to a certain threshold, and reduce benefits if earnings exceed the threshold. Therefore, the studies that focus on DI outflow investigate the effects of changing the earnings threshold or the effects of reducing benefits at different rates when earnings exceed the threshold. With respect to the earnings threshold, in Canada, [Campolieti and Riddell \(2012\)](#) study the effect of changing the amount of wages that DI beneficiaries are allowed to earn without losing DI benefits. They find a substantial labor participation increase of 5 to 6 percentage points (pp) for men and 8 to 10 pp for women due to the introduction of the earnings allowance. [Vall Castelló \(2017\)](#) analyze a reform in Spain on income tax exemption for partially disabled beneficiaries who did not work. The reform eliminated the tax exemption for those aged 55 or younger but kept it for older workers. The reform increased labor participation of the younger group by 6.5 pp. In Austria, if earnings exceed a threshold amount, DI benefits are reduced by up to 50%. [Ruh and Staubli \(2019\)](#) show substantial and sharp bunching at an earnings threshold, and that the average DI recipient with earnings just below the threshold would increase earnings by 45% in the threshold was eliminated. In Canada, [Zaresani \(2020\)](#) exploits a policy reform that changed the earnings threshold to increase labor participation. [Zaresani](#) also finds substantial bunching, which remained persistent after changing the earnings threshold, suggesting substantial adjustment costs for disabled workers.

With respect to reducing benefits when earnings exceed a threshold amount, in the US, [Weathers and Hemmeter \(2011\)](#) study a pilot project that replaced the complete loss of DI benefits due to earning above a threshold amount with a gradual reduction of DI benefits by

50% of the amount earned above an earnings disregard level. They find that the number of DI recipients with earnings above the threshold amount increases by 25% after the policy change. [Kostøl and Mogstad \(2014\)](#) analyze the consequences of providing financial incentives to DI recipients to stimulate work resumption. In 2005, the Norwegian government introduced a program where DI benefits are reduced by 60% of the earnings accumulated above an earnings threshold. They find a substantial positive effect on labor participation.

In the Netherlands, when admitted to the DI scheme, partially disabled individuals start receiving a DI benefit that also incorporates in it the unemployment insurance (UI) benefit. The benefit is reduced at a rate of 70% if beneficiaries earn wages. The duration of the benefit depends on the employment history and expires after 3 to up to 38 months from the first day of DI receipt. After this period individuals face two financial incentives to increase labor supply. If their earnings exceed a threshold amount, the benefit increases from a percentage of the minimum wage to a percentage of the pre-sickness wage, so that an increase in earnings leads to an increase in the benefit, providing a financial incentive to resume working. In addition, the UI component of the benefit expires and DI recipients can no longer claim UI. To compensate for lost benefits, individuals may increase labor supply although they may also turn to alternative social support programs.

Studying the impact of the financial incentives of the Dutch DI scheme for partially disabled individuals, [Koning and van Sonsbeek \(2017\)](#) find an average labor participation increase of 1.4 and 2.6 pp in the short- and long-term, respectively. They find a limited impact on wages. The only other study that investigates the impact of rewarding work among DI recipients is [Bütler et al. \(2015\)](#). In Switzerland, in a pilot project, a conditional cash program offered a lump-sum payment of DI benefits if beneficiaries take up or expand work so that the earnings increase is large enough to trigger a reduction in DI benefits by at least one quarter. Responses to the offer were modest: by the end of the three-year program, only 0.5% took up the offer.

Like [Koning and van Sonsbeek \(2017\)](#), we evaluate the impact of the financial incentives for work resumption. We take a local randomization approach to regression discontinuity to identify causal effects, and show that alternative identification strategies lead to similar estimated effects of the financial incentives. We contribute to the literature on the impact of financial incentives on work resumption with several findings. First, we show that, on average, financial incentives have an immediate effect of about 2.5 pp on labor participation in the first month individuals face the incentives. Accounting for anticipation and adaptation effects, the average impact increases to about 6 pp. Impact of the financial incentives is persistent over time as DI recipients continue to face the incentives.

Second, by the design of the DI scheme, the financial incentives are stronger for DI recipients who earn higher wages before they fall sick. This implies that the negative income effect of the financial incentives on labor supply is smaller for lower wage earners. DI benefits are already near the minimum wage for a large part of this group, and possibilities for increasing their incentives for work are limited. The opportunities to increase their labor participation may be limited not only by their disability but also lack of skills. We show that financial incentives has a wage gradient: Partially disabled individuals who earn at most the minimum wage before falling sick increase labor participation by 3.7 pp, while those who earn at least 2.5 times of the minimum wage increase labor participation by 7.4 pp.

Third, earlier studies provide evidence that DI reforms that limit DI eligibility can induce both work resumption and participation in alternative benefit programs ([Karlström et al., 2008](#); [Staubli, 2011](#); [Borghans et al., 2014](#); [Kantarci et al., 2023](#)). We contribute to this literature with evidence from a change in incentives built in a DI scheme instead of a change in incentives due to a DI reform. We show that, when subjected to financial incentives, partially disabled individuals do not only increase labor participation, but they also more often rely on income

from social assistance.

Fourth, earlier studies analyze the effectiveness of different policy measures to stimulate work resumption (Van Sonsbeek and Gradus, 2013; Halla et al., 2020; Kantarci et al., 2023). Different policy measures typically apply at different points in time during sickness where time spent in sickness differ. We show that financial incentives are less effective for people who spend more time in DI. Future studies should therefore account for the possibility that a certain policy measure may be ineffective because the measure does not apply at the right time during sickness, not because the particular incentive of the measure is ineffective.

Fifth, the impact of the financial incentives exhibits a strong calendar time trend partly due to the business cycle. In line with this, if disabled individuals have weaker employment opportunities in the sense that they work in sectors with low vacancy rates, their responses to the financial incentives are significantly weaker. These findings suggest that policy evaluations of reintegration measures should account for the macroeconomic environment surrounding disabled workers as one of the most vulnerable labor market groups susceptible to macro-level shocks.

Finally, we find substantial unused work capacity among temporary contract workers and the unemployed at the onset of the financial incentives. Financial incentives therefore increase work resumption much more often in these groups than in the group of individuals with permanent work contracts. Burkhauser et al. (2014) point to the growing recognition that even people with severe impairments can work up to some extent. Our finding with respect to employment status suggests that unused work capacity may also be related to employment characteristics instead of health characteristics.

The remainder of this paper is structured as follows. Section 2 describes the Dutch DI scheme for partially disabled individuals and the financial incentives. Section 3 introduces the data. Section 4 presents descriptive statistics and evidence on the impact of the financial incentives. Section 5 describes the identification strategy. Section 6 presents the estimation results. Section 7 checks the identifying assumptions. Section 8 concludes.

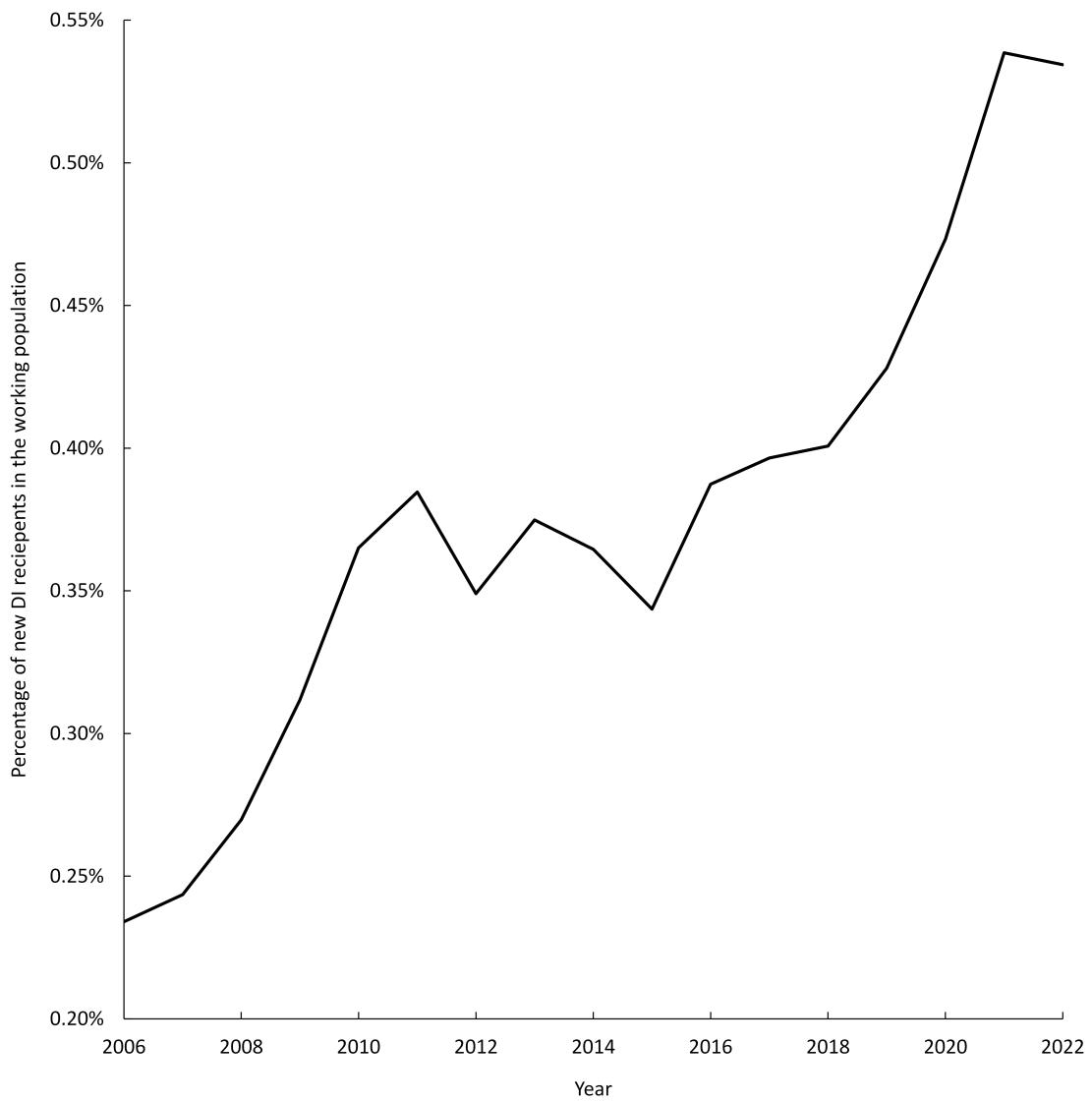


Figure 1: New DI awards for partially and temporarily fully disabled individuals in the labor force; years 2006 to 2022. Source: Statistics Netherlands.

## 2 The Dutch DI scheme and the financial incentives for work resumption

The Work and Income According to Labor Capacity Act (WIA) came into effect on 1 January 2006 for people who fell ill from 1 January 2004 onwards. In the WIA, workers who lose any part of their earning capacity due to a health impairment are entitled to the sickness benefit from their employer for a period of two years. When the sickness benefit expires, the worker can apply for a DI benefit. If the wage loss is more than 80%, with no possibility of recovery, the worker is admitted to the Income Provision Scheme for Fully Occupationally Disabled People (IVA). We do not study the labor supply responses of IVA beneficiaries, since their remaining earning capacity is very limited and there is hardly any possibility of work resumption. If the wage loss is more than 35% but less than 80%, or if the wage loss is more than 80% but there is a possibility of recovery, the worker is admitted to the Return to Work Scheme for the Partially Disabled (WGA). In the current study we focus on this scheme.

The WGA consists of two stages. In the first stage the worker is entitled to the “wage-related benefit”. This benefit consists of two components. The first component is fixed and equal to 70% of the pre-sickness (monthly) wage multiplied by the disability grade.<sup>1,2</sup> The second component is variable and equal to 70% of the pre-sickness wage multiplied by (1 - disability grade) and (1 - utilization rate). It represents remaining earning capacity that is not utilized. It is the unemployment insurance component of the DI benefit that compensates individuals who are not able to utilize their remaining earning capacity.<sup>3</sup> Duration of the wage-related benefit depends on the number of contribution years to social insurance with each year awarding one month of duration. Maximum duration is 38 months.<sup>4</sup>

When the wage-related benefit expires, individuals are entitled to one of two types of benefits depending on whether they utilize at least 50% of their remaining earning capacity. If they utilize at least 50% of remaining earning capacity, they are entitled to the “wage-supplement benefit”, which pays 70% of the pre-sickness wage multiplied by the disability grade. If they utilize less than 50% of remaining earning capacity, they are entitled to the “follow-up benefit”, which pays 70% of the pre-sickness wage multiplied by the disability degree, but the pre-sickness wage is capped at the minimum wage. Both benefits make flat-rate payments since they disregard how much individuals work once they are below or above the threshold utilization rate of remaining earning capacity. Both benefits are paid as long as the individual is disabled but expire when the individual becomes entitled to the state pension at the statutory retirement age.

These benefit rules imply two financial incentives to increase work effort when the first-stage wage-related benefit expires. The first incentive is due to the reduction in the amount of the DI benefit since the UI benefit, a component of the first-stage wage-related benefit, expires, and is no longer a component of the follow-up or the wage-supplement benefit in the second stage of the DI scheme. The second incentive is due to the reduction in the DI benefit, from the wage-supplement benefit to the follow-up benefit, if individuals do not utilize at least 50% of their remaining work capacity in the second stage of the scheme. Both incentives imply a negative income effect on labor supply.

---

<sup>1</sup>That is, the amount is fixed as long as the disability grade is not reassessed.

<sup>2</sup>During the first two months of disability the fractions are 75% instead of 70%.

<sup>3</sup>It is possible for the utilization rate to be larger than one if remaining earning capacity is incorrectly estimated. In this case, the second part of the benefit is negative. If this situation persists, it may lead to a reassessment of the disability grade.

<sup>4</sup>Beneficiaries of a DI benefit receive decision letters from the Employee Insurance Agency on two occasions. First, they receive a letter on their application for the wage-related benefit at the end of their participation in the SI scheme. Second, they receive a letter approximately three months before their wage-related benefit expires informing them of the entitlement decision of a second-stage benefit.

Figure 2 illustrates the two financial incentives for an individual who earns €100 before falling sick and has disability grade of 50%. It shows how the wage, a DI benefit, and the sum of the two change with the amount of work. Amount of work is stated in terms of the daily wage during disability as a fraction of the pre-sickness daily wage. During disability, an earned daily wage that is 25% of the pre-sickness daily wage means a remaining work capacity utilization rate of 50% (remaining work capacity is 50% given that the assumed disability grade is 50%). The figure distinguishes between two cases. In one the individual utilizes less than 50% of her remaining work capacity in the second stage of the DI scheme, and in the other she utilizes more than 50% of her remaining work capacity.

There are three notable patterns in Figure 2. First, a second-stage DI benefit is always smaller than the first-stage DI benefit. The difference is smaller the more the individual works, and it vanishes if the individual utilizes her remaining work capacity to the full extent in which case the UI component of the first-stage benefit is 0 (where red line intersects orange line). This demonstrates the first financial incentive due to the UI component of the first-stage DI benefit. Second, a second-stage DI benefit is smaller if the individual utilizes less than 50% of her remaining work capacity (and qualify for the follow-up benefit) than if she utilizes at least 50% of her remaining work capacity (and qualify for the wage-supplement benefit). This demonstrates the second financial incentive due to the 50% threshold remaining work capacity utilization rate. In Figure 2, a comparison of the total income from wages and the first-stage DI benefit with the total income from wages and a second-stage DI benefit shows the extent of the financial incentive to work in the second stage of the DI scheme when the individual utilizes less and more than 50% of her remaining work capacity. Third, individuals who earn higher pre-sickness wages face larger incentives to work because their benefit decreases by a larger amount from the first to the second stage of the DI scheme. This is illustrated by comparing Figure 2 to Figure 18 in the Appendix which shows that the financial incentives are smaller when pre-sickness wage is smaller.

The financial incentives for individuals with high pre-sickness wages are not only larger than the incentives for those with low pre-sickness wages in absolute terms, but also relative to pre-sickness wages. Although the loss of the UI component of the wage-related benefit is proportional to the pre-sickness wage, the penalty for not meeting the threshold remaining earning capacity utilization rate of 50% is based on the difference between the pre-sickness wage and pre-sickness wage capped by the minimum wage. For pre-sickness wages below the minimum wage this difference is non-existent. Above that, the difference increases with the pre-sickness wage.

The financial incentives may depend on whether the individual is entitled to the social minimum supplement. For some individuals the follow-up benefit is so low that they may qualify for the supplement, which compensates part of the reduction in benefit amounts.<sup>5</sup> Ceteris paribus, this case is more likely to apply to individuals with low pre-sickness wages, reducing their incentives even more. What is more, for individuals for whom the social minimum is higher than 70% of their pre-sickness wage, all benefits are below the social minimum and therefore there is no difference in income between being unemployed in the first and second stages of the WGA.

---

<sup>5</sup>If an individual has a partner with an income, they may get less or no social minimum supplement. Thus not everyone with very low (own) income qualifies for the supplement.

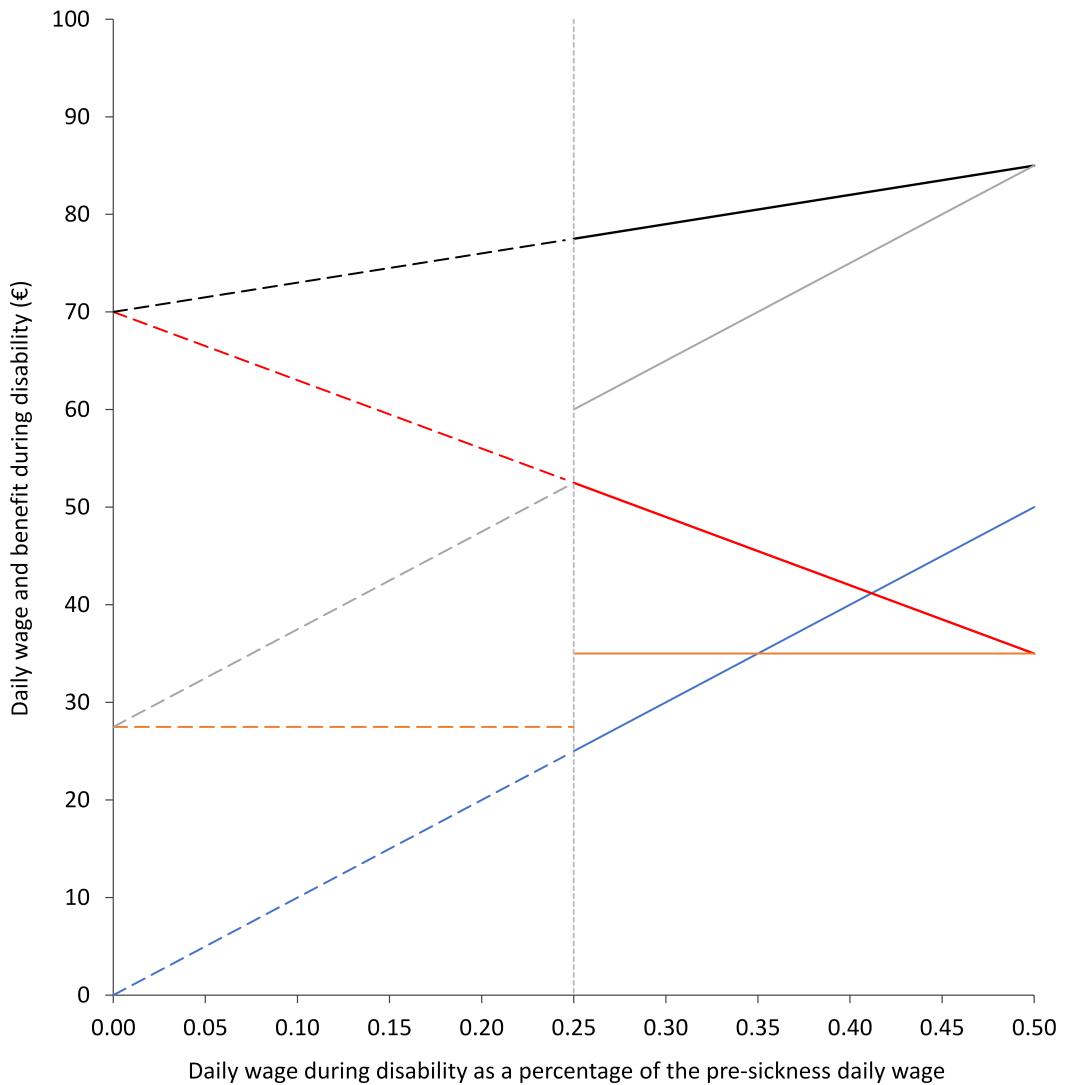


Figure 2: Financial incentives to work in the second stage of the DI scheme for an individual who earns €100 before falling sick and has a disability grade of 50%. Blue: Wage. Red: First-stage (wage-related) benefit. Orange: A second-stage (follow-up or wage supplement) benefit. Black: Sum of wage and the first-stage benefit. Gray: Sum of wage and a second-stage benefit. Vertical reference line: Remaining work capacity utilization rate = 50%. Dashed lines apply when remaining work capacity utilization rate < 50%. Solid lines apply when remaining work capacity utilization rate  $\geq$  50%. The financial incentive is the difference between the dashed black line (sum of wage and the wage-related benefit) and dashed gray line (sum of wage and the follow-up benefit) if remaining work capacity utilization rate < 50% in the second stage of the DI scheme (left side of vertical reference line). The financial incentive is the difference between the solid black line (sum of wage and the wage-related benefit) and the solid gray line (sum of wage and the wage-supplement benefit) if remaining work capacity utilization rate  $\geq$  50% in the second stage of the DI scheme (right side of vertical reference line).

### 3 Data

We use administrative data from the Employee Insurance Agency on all individuals who participated in the DI scheme for partially disabled workers (WGA). The observation period starts in January 2006, when sick individuals could apply for DI benefits the first time in the new DI scheme (WIA), and ends in December 2020. We observe their date of birth, gender, and, on a monthly basis, type of their DI benefit (wage-related, wage-supplement, and follow-up benefits), disability grade, and wages earned before reporting sick and during DI participation. We also observe whether they received unemployment insurance at the time they reported sick, and for wage earners, whether they had a permanent contract, a temporary contract, or a contract through a temporary work agency. We link these individuals to administrative data on themselves from Statistics Netherlands (CBS), with monthly information on wages and social assistance from January 2006 to December 2022, to track their labor participation status after they leave the DI scheme and hence the DI data, and to analyze participation in social assistance. Based on the available data on wages and social assistance, we define two outcome variables: dummies for labor participation and social assistance receipt based on earning positive wages and receiving positive benefits from social assistance.

The initial data set has 516,456 individuals participating in the DI scheme. To select the estimation sample, we drop individuals who re-enter the DI scheme after exiting, or those who move back to the first stage of the DI scheme from the second stage. Furthermore, we drop individuals if they participate in the DI scheme for fully and permanently disabled (IVA) since the rules and incentives for them are quite different. Finally, we drop individuals if they exit the DI scheme during the first stage of the scheme, or do not reach the second stage of the scheme by the end of the observation period, since these individuals do not experience the financial incentives to resume work.

### 4 Descriptive statistics and evidence on the impact of financial incentives

Table 1 presents sample means of some background characteristics for three groups based on pre-sickness wages earned as a fraction of the minimum wage, as well as outcomes in the last month of the first stage of the DI scheme before individuals face the financial incentives, and in the first month of the second stage of the DI scheme when they first face the financial incentives. Individuals with higher pre-sickness wages are older and much more often male. They are more likely to have a disability grade larger than 50% but this is due to the correlation between disability grade and pre-sickness wages earned.<sup>6</sup> Higher pre-sickness wage earners appear to have a longer work history and therefore receive the first-stage wage-related benefit for a longer period (Section 2). They receive the second-stage wage-supplement benefit for a longer period and the follow-up benefit for a shorter period.<sup>7</sup> It might be that due to better job opportunities or stronger labor market attachment, they are more often able to utilize more than 50% of their

<sup>6</sup>Disability grade is defined as the difference between reference wage and assessed remaining earning capacity, divided by the reference wage. The reference wage is the theoretical wage of someone who is not disabled but similar to the partially disabled individual in all characteristics. The remaining earning capacity is based on a set of jobs that the partially disabled person can still do. For individuals with low pre-sickness wages –and therefore likely low reference wages– there will not be jobs paying hourly wages far smaller than their pre-sickness wages since the lowest they can earn is the minimum wage. For individuals with higher pre-sickness wages, however, remaining earning capacity can be considerably below their reference wage. Therefore, disability grade can have a strong correlation with pre-sickness wage.

<sup>7</sup>Figure 19 in the Appendix shows the distribution of the number of months spent in the first and second stages of the DI scheme.

remaining earning capacity to become entitled to the wage-supplement benefit that is higher than the follow-up benefit. This is in line with their considerably higher labor participation rate, and smaller social assistance receipt rate at the bottom of Table 1. For each wage group, Table 1 shows the difference in labor participation from the last month of the first stage to the first month of the second stage of the DI scheme. The difference is positive, providing first evidence on the impact of financial incentives.

Figure 3 shows the sample average of labor participation during two periods of two years, before and after the financial incentives of the DI scheme take effect when the first stage of the DI scheme expires. We distinguish among five wage groups. Labor participation remains stable or slightly increases for the three highest wage groups as they exhaust their first-stage benefit. For lower wage groups, however, labor participation shows a notable decrease. This suggests that low-wage earners struggle to find suitable jobs or recover from disability to utilize their remaining work capacity as they spend time and exhaust their first-stage benefit. Labor participation of all wage groups increase prior to and after the cutoff month when financial incentives take effect, suggesting anticipation and adaptation effects of the financial incentives. Compared to the first stage, higher wage groups show notably higher labor participation rates during the second stage of the DI scheme where they face stronger financial incentives to resume work than do lower wage groups. Figure 4 shows the sample average of social assistance receipt. The lowest wage group is much more likely to rely on social assistance since they more often become entitled to social assistance due to earning low wages. For all wage groups, however, social assistance receipt shows a discontinuous increase when the first stage of the DI scheme expires. Apparently, in all wage groups, some individuals struggle to increase their labor participation or earnings after they exhaust their first-stage benefit, and substitute labor income with income from social assistance.

In Section 2 we demonstrated how financial incentives could generate a negative income effect to induce work resumption. The negative income effect is particularly large if partially disabled individuals work in the first stage but do not work in the second stage of the DI scheme while they face the financial incentives to work: vertical distance between any point on the (dashed or solid) black line and where the dashed gray line intersects the y-axis in Figure 2. Facing this financial incentive, forward-looking partially disabled individuals at the end of the first stage of the DI scheme are expected to consider resuming work in the second stage of the DI scheme. To check this empirically, for each individual in the data, we calculate a potential financial incentive measure, and analyze how it correlates with the work resumption decision. We define the potential financial incentive measure as the ratio of the income from wage and wage-related benefit if individuals work (as observed) in the first stage of the DI scheme to the income from follow-up benefit if individuals do not work (as assumed) in the second stage.<sup>8</sup> Figure 5 plots the work resumption decision –change in labor participation from the last month of the first stage to the first month of the second stage of the DI scheme– against the potential incentive measure, where we condition labor participation on a set of observable characteristics.<sup>9</sup>

---

<sup>8</sup>The income in the first stage depends on the pre-sickness wage, disability grade and how much of the remaining earning capacity is utilized (Section 2). Pre-sickness wage is given. We consider remaining earning capacity utilization rate as the ratio of wage income while disabled to remaining earning capacity (difference between pre-sickness wage and wage while disabled). We consider wage income as observed at event month -5, i.e. 5 months before the wage-related benefit expires. The income in the second stage depends on the pre-sickness wage and disability grade. We assume that disability grade remains constant in the first and second stages of the DI scheme.

<sup>9</sup>In particular, labor participation represents the residuals from a regression of labor participation on observable characteristics including labor participation status 5 months before the financial incentives take effect, age, gender, sector, employment status at the time of applying for DI (had a permanent job, temporary job, unemployed), quadratic function of average time spent in the first stage of the DI scheme, and average of the business cycle indicator over a period of 2 years after the financial incentives take effect.

Work resumption is strongly positively related to the potential financial incentive. For example, a value of 2 for the measure, representing a 50% decrease in income due to stopping work when the first stage of the DI scheme expires, is associated with a labor participation increase of about 7 pp on average. In the next section we investigate if the impact of the financial incentives on work resumption is causal.

Table 1: Sample means of background characteristics and labor market outcomes by pre-sickness wage groups in months before and after financial incentives take effect

	Pre-sickness wage as a fraction of the minimum wage						
	<100%			150–200%		>250%	
	Before	After	Difference	Before	After	Difference	Before
Age	46.163	46.247	0.083	48.734	48.817	0.083	53.605
Male	0.203			0.508			0.721
Permanent contract	0.407			0.544			0.554
Temporary contract	0.289			0.214			0.182
Unemployed	0.270			0.238			0.264
Disability grade >50%	0.203	0.204	0.001	0.510	0.512	0.002	0.455
Months receiving wage-related benefit	19.607			22.306			25.541
Months receiving wage-supplement benefit	12.803			21.326			22.271
Months receiving follow-up benefit	27.932			22.976			18.814
Pre-sickness daily wage (€)	48.290	48.296	0.006	102.783	102.789	0.006	184.370
Labor participation	0.362	0.388	0.026	0.452	0.479	0.027	0.512
Social assistance participation	0.046	0.048	0.002	0.000	0.003	0.003	0.000
Number of individuals	2,496			12,318			10,143

Notes: Labor participation is defined as having positive wage income. Pre-sickness daily wage and daily wage while disabled are in January 2006 euros (correcting for indexation of wages with respect to price and sectoral wage inflation).

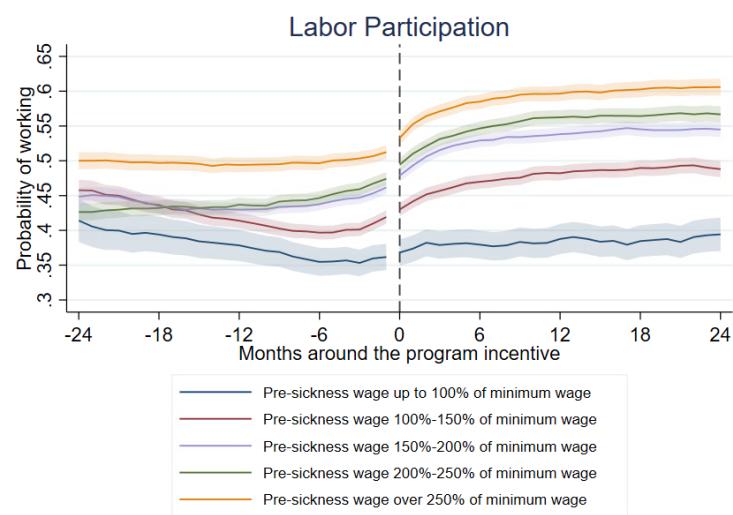


Figure 3: Labor participation around the month financial incentives take effect.

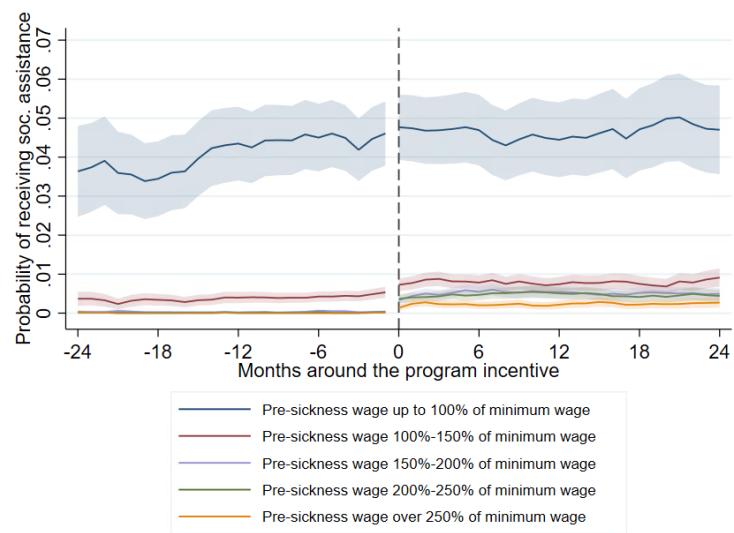


Figure 4: Social assistance receipt around the month financial incentives take effect.

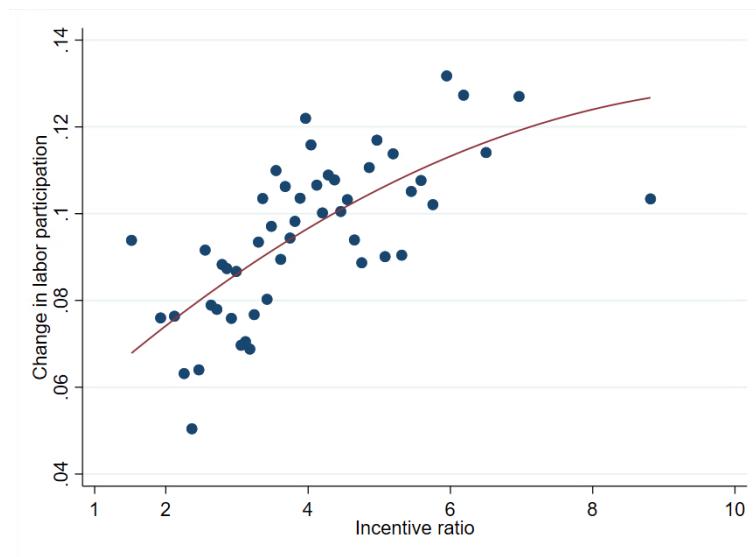


Figure 5: Potential financial incentive and work resumption decision.

## 5 Identification strategy

As described in Section 2, individuals with higher pre-sickness wages face a stronger financial incentive to resume work when their first-stage DI benefit expires. Therefore, we distinguish across five pre-sickness wage groups to analyze if the program leads to higher labor participation responses among those with higher pre-sickness wages.

The estimation strategy to identify the effects of the financial incentives is based on a Regression Discontinuity (RD) design. This design requires a cutoff in the running variable where the probability of treatment changes discontinuously. Our running variable is the time around the moment individuals move from the first to the second stage of the DI scheme. It is defined such that it equals 0 in the first (event) month the individual receives a second-stage benefit (wage-supplement or follow-up), 1 in the month after, -1 in the month before, etc. By this construction, probability of treatment is 0 for all negative values of the running variable, and 1 for all nonnegative values. Since the running variable takes on only integer values, the cutoff can in principle be placed anywhere between  $-1$  and  $0$  without affecting the treatment assignment for any of the observations. This gives rise to a sharp RD design as compliance is perfect.

As the running variable takes on discrete values, use of the standard continuity-based RD approach is questionable, especially given the fact that the number of distinct values of the running variable is as small as 234. That is, when the running variable is discrete and a continuity-based RD design is applied, the effective number of observations used is the number of the distinct values, not the total number of observations (Cattaneo et al., 2020). For our study, this means estimating the treatment effect using 234 distinct values of the running variable, instead of the available 3,365,133 monthly observations for 46,250 individuals. In this case Cattaneo et al. explain that the Local Randomization (LR) approach to RD may be the only valid RD approach. This approach requires stronger assumptions than the continuity-based RD approach. However, if its assumptions are met, it allows the running variable to be discrete and to identify still a causal effect.

The rationale is that there exists a window around the cutoff where the data behaves as if it were part of an experimental setup in which units receive a random score (value of the running variable), and are assigned to treatment if and only if their score is equal to or above the cutoff value. Consequently, the value of the running variable is not related to any observed or unobserved characteristic of the units, other than by chance. As long as the assigned score by itself does not affect the outcome, the expected value of the outcome variable, conditional on treatment status, is then the same for all values of the running variable within the window, even if the control and treatment groups are only observed in (mutually exclusive) parts of the window. This is illustrated in Figure 6. The difference between the expected values of the control and treatment groups identifies the treatment effect. This difference can be estimated by the difference in the sample averages of the two groups.

The LR approach to RD identifies the treatment effect around the cutoff. As we do not take a continuity-based approach to RD, we do not make continuity assumptions to identify and estimate a causal effect at the cutoff (Mattei and Mealli, 2017). Furthermore, Cattaneo et al. (2015), Mattei and Mealli and Sales and Hansen (2020) propose different sets of assumptions within a neighborhood of the cutoff to interpret a RD design as a local randomized experiment. We adapt the approach of Cattaneo et al..

### 5.1 Identification assumptions

Let  $r_0$  be the cutoff value of the running variable,  $T$  the treatment indicator, and  $W_0$  the local randomization window. The potential outcomes for hypothetical values of the running variable

and treatment status of individual  $i$  are denoted with  $Y_i(r, t)$ . There are two assumptions that need to hold for the validity of the LR approach:

1. Unconfounded assignment. The distribution function of the running variable inside the window,  $F_{R_i|R_i \in W_0}(r)$ , does not depend on the potential outcomes and is the same for all units:  $F_{R_i|R_i \in W_0}(r) = F_0(r)$ , where  $F_0$  is any distribution function.<sup>10</sup>
2. Exclusion restriction. The potential outcomes do not depend on the value of the running variable inside the window, except via the treatment assignment indicator:  $y_i(r, t) = y_i(t) \forall i$  such that  $R_i \in W_0$ .

The first identifying assumption has two requirements. First, it requires that potential outcomes do not affect the running variable inside the window. This implies that the running variable is not endogenous to potential outcomes. Second, it requires that distribution of the running variable is the same for all individuals within the window.

Our running variable is determined as follows. For each individual, it starts at a certain negative value depending on the individual's work history, increases incrementally over time until the individual exploits the first-stage benefit, and it remains increasing in the second-stage until the individual exits the scheme or until it is the end of the observation period.<sup>11</sup>

With respect to the first requirement, we distinguish between the first- and second-stage durations of the DI scheme that span the values of our running variable. Since the duration of the first-stage benefit is institutionally determined, it cannot be adapted in response to the changes in labor supply, or to factors correlated with the labor supply of beneficiaries. Therefore, neither the first-stage duration nor the treatment assignment at the end of the first stage can be endogenous. As we control for individual-specific fixed effects, individual attributes that are time invariant and affect the running variable in the second stage through labor supply cannot confound the treatment effect estimate. For example, individuals with strong relationships with their employers could have better opportunities to resume work when they face the financial incentives of the program in the second stage, and hence have a shorter second-stage duration. However, as long as this is a time invariant attribute, our estimate of the treatment effect will not be biased.

The second requirement of first assumption is also likely to be met. Individuals cannot manipulate the value of the running variable when they are in the DI scheme. Furthermore, due to the panel nature of the available data, we have observations on (mostly) the same people before and after the cutoff. Therefore, people on both sides of the cutoff should have similar (unobserved) characteristics.

There is, however, a caveat regarding the timing of entry into and exit from the DI scheme. If this was the same for all individuals, they would all have experienced same values of the running variable, regardless of their observed (e.g. duration of the first-stage benefit) or unobserved characteristics. In particular, this would imply that their potential outcomes cannot affect the running variable. In fact, this holds for the smallest possible window,  $\{-1, 0\}$ . Everyone in the study sample experiences both the last month of the first stage and the first month of the second stage of the DI scheme. However, as the window widens, more and more individuals have missing observations within the window at one or both sides of the cutoff. These missing observations may not be distributed randomly across individuals. The duration of the first stage of the DI scheme, and thereby the number of observations below the cutoff, is determined by

---

<sup>10</sup>We do not assume that the distribution function is known. This is not needed because in the analysis we do not make use of finite-sample inference due to the large sample available (Cattaneo et al., 2020).

<sup>11</sup>Duration of the first-stage benefit is determined by the duration of unemployment insurance which depends on the work history (Section 2).

the work histories of individuals. Individuals with long work histories may systematically differ from those with short work histories. For example, individuals with higher potential outcomes may have been more likely to work before falling sick and therefore have a longer first-stage duration. Similarly, people who exit the DI scheme soon after the cutoff may be systematically different from those who do not, although a relatively small part of the sample (14.5%) exits the scheme before the end of the observation period, of whom less than half exit in the first 12 months after the cutoff. We assume that missing data is random.

The second identifying assumption implies that if there was no treatment, the expected value of the outcome should be the same for all values of the running variable inside the window  $W_0$ . Indeed, due to the exogenous variation in the duration of the first stage, and hence the moment of entry into the DI scheme in terms of calendar time, the month of the cutoff is like any other month, and therefore it is not immediately clear why being in a certain month before or after it should matter for the outcome (labor participation) in the absence of treatment.

However, the running variable does contain information about the passage of time. For each individual this variable is perfectly correlated with other variables related to time such as the time spent in the DI scheme, age, and calendar time. These variables may affect labor participation. Below we discuss how we account for these variables in the regression analysis.

As mentioned above, exploiting the panel dimension of the available data, we control for individual-specific fixed effects to account for time-invariant observed and unobserved characteristics that may affect labor participation. Therefore, the two identifying assumptions imply that, conditional on time-variant controls and time-invariant individual attributes, potential outcomes and the running variable are not related except through treatment assignment.

## 5.2 Identification of the full treatment effect

To identify the full treatment effect, we require additional assumptions related to the timing of the responses to the financial incentives. Since the model is based on a comparison of the outcome before and after the cutoff, we need that people close to the cutoff, but still in the first stage of the scheme, do not already change their labor participation in anticipation of the program incentives of the second stage. The modest increase in labor participation of all wage groups in Figure 3 might cast doubt on this assumption, as it could be explained by anticipation of the incentives.

There can be, however, alternative explanations for the different trends observed just before the cutoff. For example, changes in the outcome prior to the treatment could be due to that treatment is endogenous, instead of that they are due to anticipation (Malani and Reif, 2015). However, as argued above, this is impossible since the duration of the first stage of the DI scheme is institutionally determined.

Still, for higher pre-sickness wage groups, if part of the labor supply response already takes place before the cutoff, the treatment effect will be underestimated. A similar issue arises for the labor supply response after the cutoff. If individuals are able to utilize their remaining work capacity in suitable jobs at any given point in time, they could react to the program incentives immediately after the cutoff. In practice, however, it may be unrealistic to expect individuals to time their response so precisely. The labor supply response may be spread out farther beyond the cutoff due to adaptation. In this case the RD estimator will, again, underestimate the true treatment effect.

Underestimation of the treatment effect due to anticipation before the cutoff, or adaptation after the cutoff, is not solved by taking a wider window, although the estimates may be closer to the true effect. For the latter, for example, consider the case where the no-anticipation assumption holds, but after the cutoff it takes  $r$  months ( $r > 0$ ) before all program participants

have responded. Conditional on covariates and individual fixed effects, the estimation in the window  $\{-r - 1, \dots, r\}$  compares the mean in the  $\{0, \dots, r\}$  window to the mean in the  $\{-r - 1, \dots, -1\}$  window. However, the average effect in the  $\{0, \dots, r\}$  window will be below the full effect at  $r$ .

These considerations imply that the estimated effects of treatment can be viewed as lower bounds to the true full treatment effects. However, even if the assumptions regarding the timing of the response are violated, the differences between pre-sickness wage groups remain equally informative as long as individuals in the different groups violate the assumptions in the same way.

### 5.3 Window selection

An important step in the LR design is window selection. Ideally, the selected window is the widest possible window within which the LR assumptions hold. However, this cannot be tested directly. [Cattaneo et al. \(2020\)](#) describe two options to select the window. The first option takes a data-driven approach. The window size estimate is the window within which one or more observed covariates do not change with the running variable but can change outside it. Such a covariate does not exist in our data. This leaves us with the second option, which is choosing the window in an ad-hoc manner. We consider a window size that is symmetric around a hypothetical value of  $-0.5$ , and therefore contains the same number of months before and after the cutoff.

Several considerations should be weighed against each other when choosing the window. On the one hand, the wider the window, the more precisely coefficients can be estimated, as more observations are used. On the other hand, the smaller the window, the more likely that the LR assumptions hold. As (symmetric) windows are nested within each other, it follows that if there is some window  $W_0$  in which the LR assumptions hold, they will necessarily also hold in smaller windows. However, as the window gets wider, the less plausible it becomes that there is no relation between the running variable and the outcome variable anywhere in the window. In the context of the current study, the further away in time from the cutoff, the more likely that there are changes in (unobserved) individual characteristics that may affect labor participation, such as health or household income. If these changes generally move in a specific direction over time, for example health deteriorates over time, estimation of the treatment effect will be biased.

In terms of the assumptions related to the timing of the response (no anticipation before the cutoff and no adaptation after it), the size of the window seems irrelevant. The assumption that the effect of treatment is observed immediately after the cutoff needs to hold in every window, because every window includes the immediate term. Furthermore, it is hard to conceive of a situation in which there is an anticipation effect some time before the cutoff but not right before it. Only then would the assumption of no anticipation hold in small windows but not in wider ones.

Given the large sample sizes in all pre-sickness wage groups, the variance of the coefficient estimates is less of a concern than the potential bias introduced if the LR assumptions do not hold. Even in the smallest possible window there are some 5,000 to 25,000 observations in the data for each group. Therefore, the smallest possible window is preferred in this study, which is  $\{-1, 0\}$ . We consider alternative time windows to show sensitivity.

## 5.4 Regression specification

We estimate the following regression model:

$$y_{ir} = \alpha_i + \beta T_{ir} + X_{ir}\gamma + \varepsilon_{ir}. \quad (1)$$

$i$  indexes individuals.  $r$  indexes the event time. Its values from -60 to -1 indicate the months before individuals face the financial incentives, 0 is the cutoff month when first subjected to the program, and 1 to 92 are the remaining months of the program. The dependent variable, labor participation or social assistance receipt, is given by  $y$ . The coefficient  $\beta$  on the treatment indicator  $T$  is the main parameter of interest. Assuming that the identifying assumptions hold, it captures the mean effect of the financial incentives around the cutoff.  $X$  is the vector of covariates related to the running variable. It includes dummies controlling for multiples of six months of DI receipt duration, and a linear function of the business cycle indicator. The dummies capture the potentially strong relationship between labor participation and DI duration. The business cycle indicator captures macroeconomic shocks. We consider a linear functional form as suggested by the binned scatter plot of labor participation against the (lagged) indicator in Figure 21 in the Appendix.<sup>12,13</sup>  $\alpha_i$  is an individual-specific constant. It captures time-invariant labor participation differences across individuals. It also captures differences in pre-sickness wage levels or time-invariant health conditions.  $\varepsilon_{ir}$  is an idiosyncratic (unobserved) shock, assumed to be uncorrelated with all explanatory variables. Only observations for which the value of the running variable falls within the window,  $r \in W_0$ , are used in the estimation.

---

<sup>12</sup>The indicator is constructed by the Dutch Central Bank to identify turning points in the Dutch business cycle (Butler et al., 2019). It is composed of 86 potential sub-series that are closely related to the development of real GDP growth and are also up to six months ahead of it. Figure 20 in the Appendix shows that the indicator is reasonably capable of identifying tipping points in real GDP growth from the recent past. We allow a period of 6 months for the indicator to affect labor participation. Since the indicator itself predicts 6 months ahead, we lag the indicator by 12 months. When the indicator is lagged 12 months, Figure 21 shows a positive impact on labor participation as we would expect. It shows a negative impact when lagged 6 months.

<sup>13</sup>We cannot consider calendar time and age in the vector of covariates because they are perfectly collinear with the dummies for DI duration. We consider that the business cycle indicator captures the macroeconomic shocks that calendar time would capture. To investigate the possible effect of age on the treatment, we conduct heterogeneity analysis with respect to age.

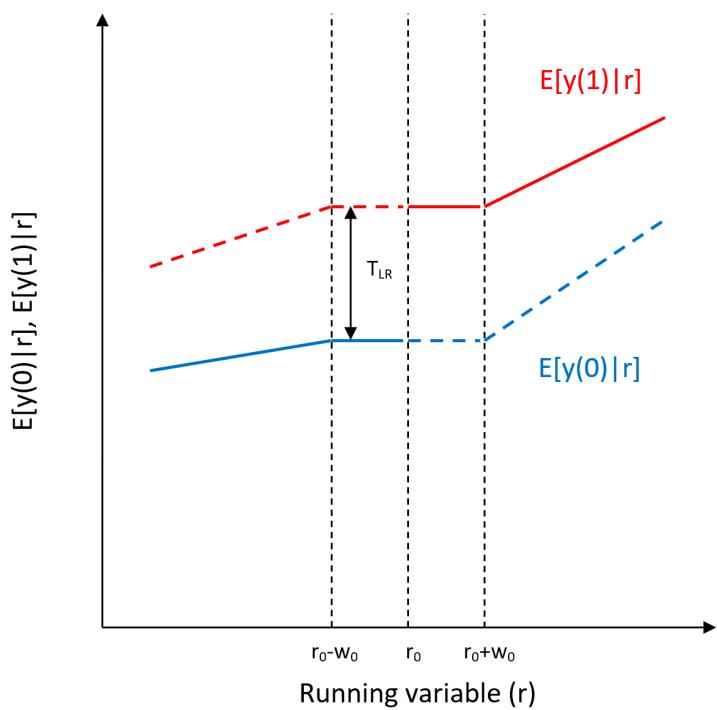


Figure 6: Illustration of the local randomization approach to RD based on [Cattaneo and Titunik \(2021\)](#).

## 6 Effect of the financial incentives on labor participation

### 6.1 Baseline effects

Table 2 presents the estimated effects of the financial incentives for five wage groups in the  $\{-1, 0\}$  window.<sup>14</sup> In line with the descriptive predictions made in Section 4, financial incentives exhibit a wage gradient: As the negative income effect of the financial incentives are stronger for higher income earners, they respond more strongly. These responses might, however, partly reflect health responses such that higher income groups may be more able to invest in health care and recover from sickness more often than lower income groups can. In Section 7, we check if a health outcome exhibits a discontinuous change at the cutoff month when financial incentives take effect or at later months.

Figure 7 presents the estimated effects of the financial incentives based on symmetric time windows around the month financial incentives take effect. We consider time windows of different sizes of up to 48 months. As for the smallest time window, financial incentives induce stronger responses among higher income groups. Besides this, the effects get larger as the time window gets wider. This is most notable for time windows of 4 to 8 months around the month financial incentives take effect. This may indicate anticipation and adaptation effects as suggested by the descriptive trends around the cutoff month in Figure 3. We explore this possibility in the next section. Magnitudes of the estimated effects from windows of sizes 10 to 24 months are similar to each other. This suggests that a large part of the impact of the financial incentives is realized within four months after individuals face the incentives. The estimated effects of the financial incentives in the long term, that is those when the time window is 24 months or wider, should be interpreted with caution. In the LR design, potential outcomes can depend on the running variable outside the window  $W_0$  (Figure 6). This is because the LR assumptions do not need to hold outside the window. This, however, also means that the LR estimates outside the window do not need to reflect causal effects. Therefore, we take the long-term effect estimates as suggestive of causal effects.

Figure 8 presents the estimated effects of the financial incentives on social assistance receipt. The effects are significant except for the lowest wage group. As suggested by Figure 4, this group receives social assistance much more often than other groups before or at the onset of facing financial incentives. They are therefore less likely to make new claims on social assistance when they face the financial incentives. The significant effects for other wage groups suggest that some individuals struggle to increase their labor participation after they exhaust their first-stage benefit, and substitute labor income with income from social assistance. This finding is in line with earlier studies that show that DI reforms that limit DI eligibility induce work resumption but also participation in alternative benefit programs to compensate for the lost DI benefits (Karlström et al., 2008; Staubli, 2011; Borghans et al., 2014; Kantarcı et al., 2023).

---

<sup>14</sup>In the regressions presented in this section, the dummies controlling for multiples of six months of DI receipt duration are almost always significant except in the smallest window and for the lowest wage group. The linear function of the business cycle indicator is almost always insignificant.

Table 2: Estimated effects of financial incentives on labor participation among pre-sickness wage groups

Pre-sickness wage as a fraction of minimum wage	Labor participation
<100%	0.008** (0.003)
100–150%	0.012*** (0.002)
150–200%	0.017*** (0.002)
200–250%	0.020*** (0.002)
>250%	0.020*** (0.002)

Notes: The window size is 2 months in each regression. Regressions are based on 4,992, 20,908, 24,636, 21,678 and 20,286 observations for the lowest and higher wage groups respectively. Regressions control for time spent receiving the wage-related benefit, (lagged) business cycle indicator, and fixed effects. \*\*\*, \*\*, \* denote statistical significance at 1%, 5% and 10%, respectively. Standard errors (in parentheses) account for heteroskedasticity and clustering at the individual level.

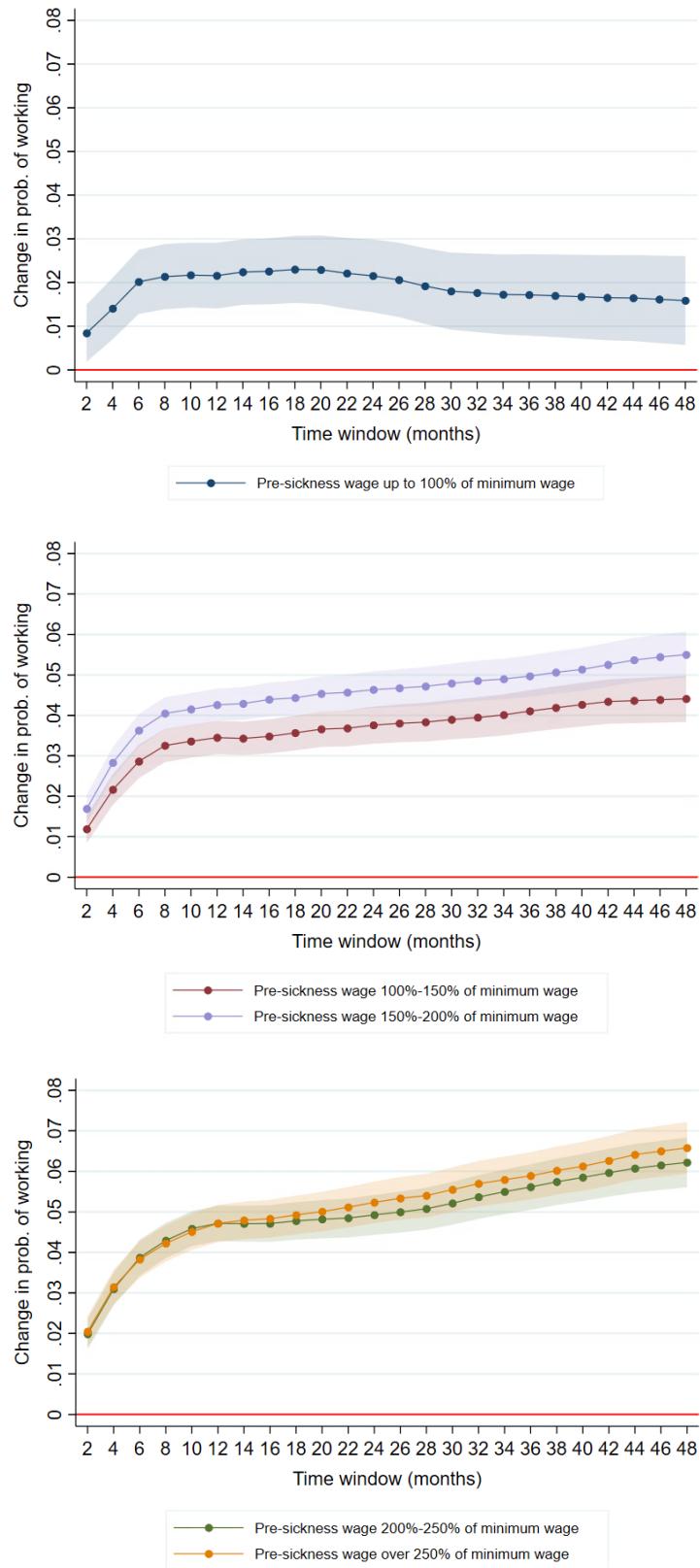


Figure 7: Estimated effects of financial incentives on labor participation and 95% confidence intervals around them for different time windows around the month financial incentives take effect.

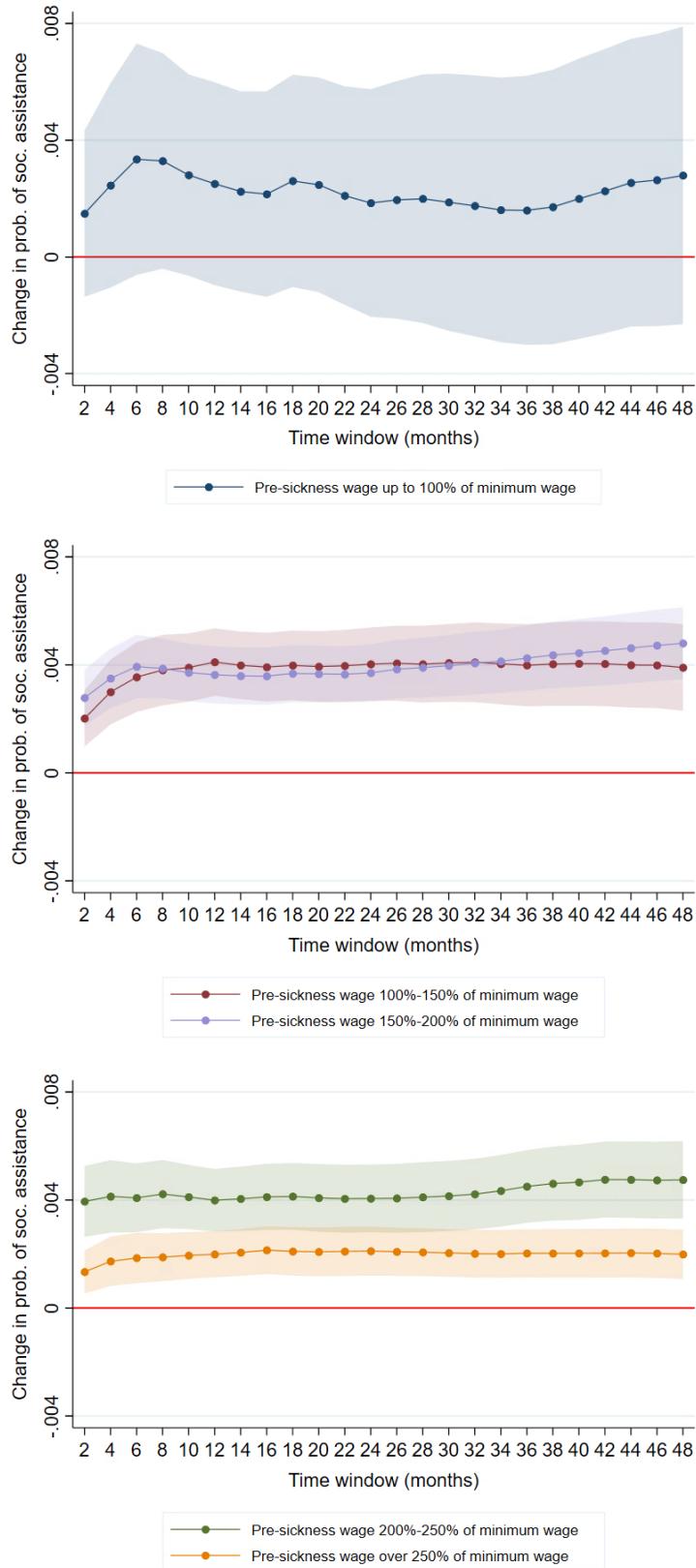


Figure 8: Estimated effects of financial incentives on social assistance receipt and 95% confidence intervals around them for different time windows around the month financial incentives take effect.

## 6.2 Full treatment effect

Identification of the true effect of the financial incentives requires that individuals do not change work decisions in anticipation of the financial incentives before they face the financial incentives, or do not take time to respond to the financial incentives due to adaptation after they face the financial incentives the first time at the cutoff (Section 5.2). In Figure 3, notable increases in labor participation close to the month individuals face the financial incentives suggested anticipation and adaptation effects. In Figure 9 we investigate transitions between working and not working in adjacent months. There is a trend of a slightly decreasing work resumption throughout the four-year period before and after individuals face the financial incentives. It might be that labor market opportunities or attachment decrease over time. There is a notable increase in work resumption in the few months preceding and succeeding the month financial incentives take effect, while transitions to not working are much less pronounced. This provides supporting evidence for anticipation and adaptation effects.

Assuming that anticipation and adaptation are part of the true responses, we investigate to which extent they lead to an underestimation of the true effect of the financial incentives in the baseline regression analysis. We estimate a “donut hole” regression where we exclude observations within the time window of 8 months around the cutoff month, keeping otherwise the baseline regression specification the same. A window of 8 months is chosen based on the observation that, in Figure 7, treatment effect estimates show a notable increase in windows of 2 to 8 months for all wage groups, possibly due to anticipation or adaptation effects. Figure 10 presents the estimation results. The treatment effect estimates are substantially larger than those based on the baseline regression. This suggests that baseline treatment effects might be underestimated due to anticipation of the financial incentives or adaptation to it. Therefore, the baseline treatment effect estimates can be viewed as lower bounds to the full treatment effect of the financial incentives.



Figure 9: Fraction of people working (not working) conditional on not working (working) in the previous month during 24 months before and after the cutoff month financial incentives take effect.

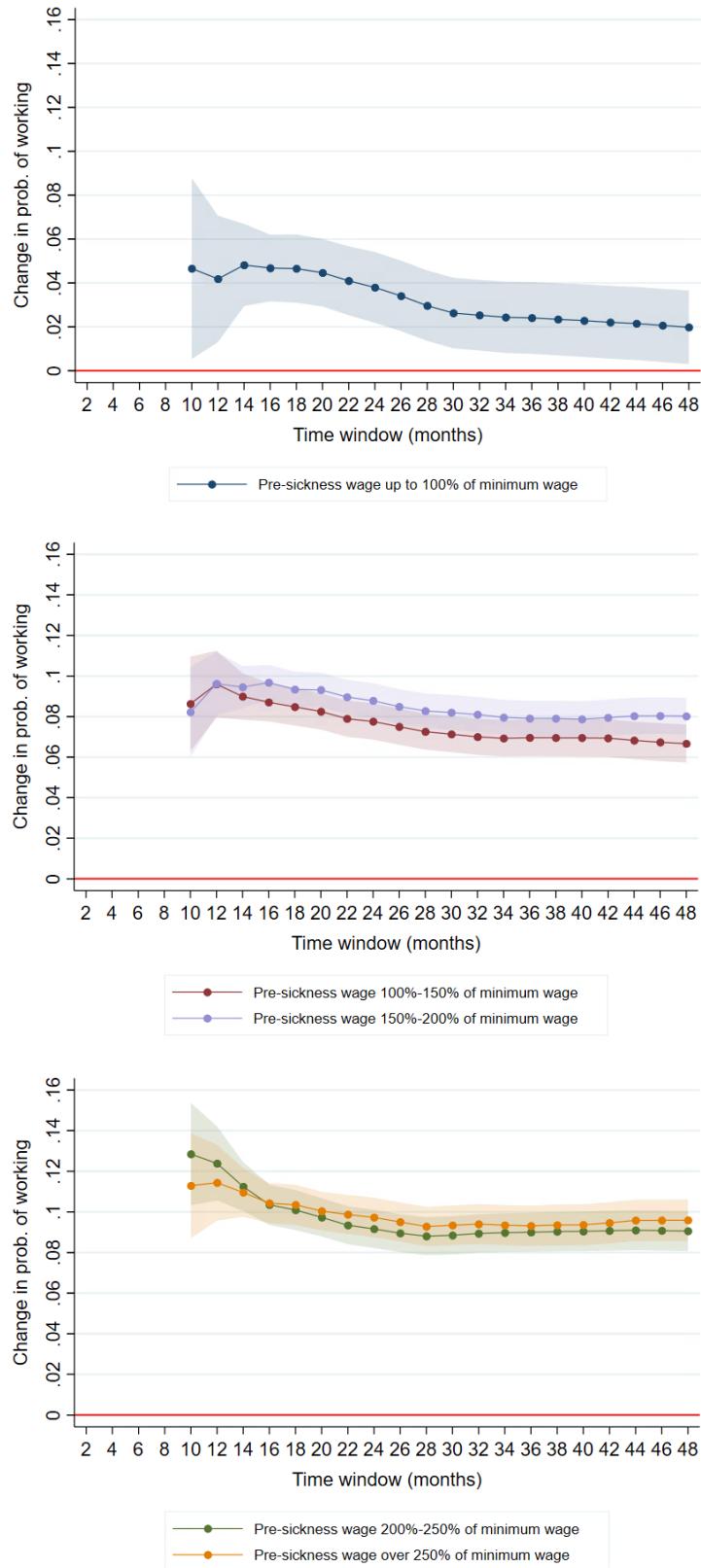


Figure 10: Estimated effects of financial incentives and 95% confidence intervals around them based on donut hole regressions.

### 6.3 Heterogeneous effects

We investigate the heterogeneous impact of financial incentives with respect to disability and labor market characteristics. We hypothesize that duration of the first-stage benefit, calendar year of entry to the DI scheme, employment status at the time of reporting sick (permanent job, temporary job, or unemployed), and the sectoral vacancy rate at the time of reporting sick are important factors of how financial incentives affect work resumption. Each and any of these four factors, however, can be correlated with other observable and unobservable characteristics, making it difficult to attribute an impact to only the factor itself. For example, individuals with a short duration of the first-stage benefit will have short work histories and hence be younger. They will then exhibit a stronger response to the financial incentives. To address this, we consider a set of covariates as main variables driving compositional differences across duration groups. We weigh individuals to have similar distributions of these covariates across duration groups. Following [Hainmueller \(2012\)](#), we generate sample weights using entropy balancing. Individuals are weighted to adjust inequalities in representation with respect to the first moment of the covariate distributions. As covariates, we consider age, gender, contract type, and year of entry to the DI scheme. We then estimate Equation (1) within each duration group using the re-weighted sample of that group.

#### Duration of DI receipt

A large number of studies analyze which policy measures are effective to limit inflow into DI or stimulate outflow from it. It is argued that measures that tighten criteria to enter DI are more effective than the measures that stimulate reintegration among DI recipients ([Van Sonsbeek and Gradus, 2013](#); [Haller et al., 2020](#); [Kantarcı et al., 2023](#)). These studies do not account for the time spent in sickness. A longer sickness period may lead to more human capital loss or a stronger scarring effect, reducing the prospects of finding a job ([Arulampalam, 2001](#); [Arulampalam et al., 2001](#)). Therefore, labor market attachment or opportunities may be very different when people apply for DI the first time and when they are already receiving DI. Same people may then respond to a same policy measure differently at different points in time during sickness where time spent in sickness differ. An implication of this is that, as different policy measures typically apply at different points in time during sickness, it becomes difficult to identify and compare the effectiveness of the idiosyncratic incentives of these measures to stimulate work resumption.

In our data, we observe complete episodes of DI receipt unless individuals continue to claim DI beyond the observation period. Figure 19 in the Appendix presented the distribution of the number of months individuals claim DI in the first and second stages of the DI scheme. Exploiting the random variation in the number of months individuals spend to exhaust their first-stage benefit and face the financial incentives, we analyze how the effectiveness of financial incentives as a policy measure to induce work resumption depend on the amount of time people spend in DI.

Figure 11 presents the estimated effects of the financial incentives for two groups of individuals who have spent at most and at least 2 years in the first stage of the DI scheme. The responses of the individuals with longer durations of DI claiming are much weaker than the responses of individuals with shorter durations. This suggests that financial incentives to stimulate work resumption becomes less effective the more time individuals spend in the DI scheme.

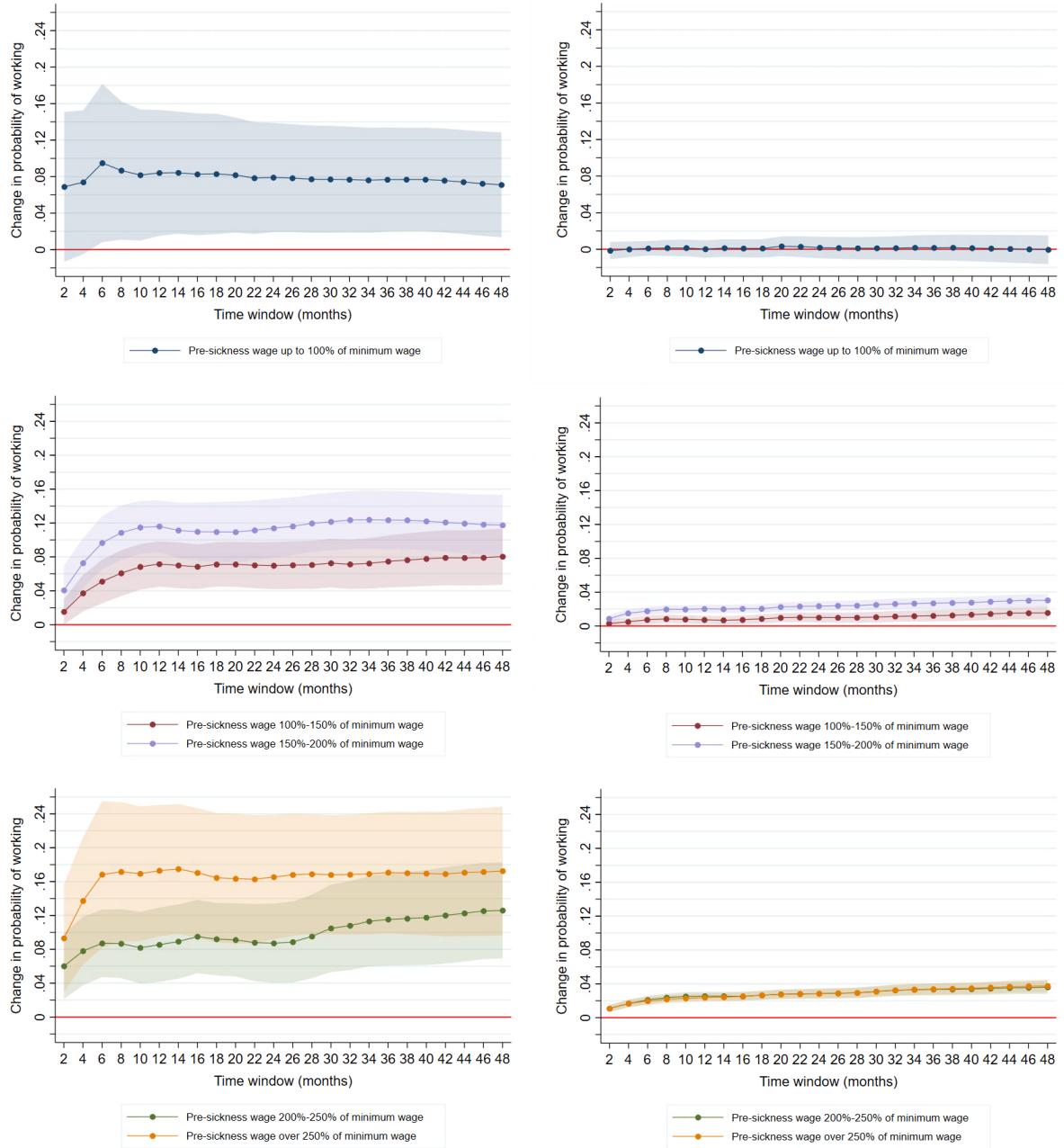


Figure 11: Estimated effects of financial incentives and 95% confidence intervals around them for individuals who spend at most (left panel) and at least (right panel) 2 years in the first stage of the DI scheme. The group with 2 years of duration is re-weighted so that the duration groups have similar distributions of age, gender, contract type, and year of entry to the DI scheme.

## Year of entry to the DI scheme

Impact of the financial incentives of the DI scheme may change over calendar time for different reasons. Employment opportunities may depend on the business cycle and affect how disabled individuals respond to the financial incentives to resume work. During downturns of the business cycle, employers may find employees with disabilities less attractive, whereas employees with disabilities may try harder to find suitable jobs. The impact of the financial incentives might also change over time due to factors endogenous to how beneficiaries manage financial incentives. Over time, financial literacy of DI awardees may improve and awardees of later years may have a better understanding of the implications of the financial incentives for their income and respond more strongly. We analyze the impact of the financial incentives across cohorts with respect to the year they were awarded DI benefits. Figure 12 presents the estimated effects of financial incentives for three-year cohorts who were awarded DI during the observation period from 2006 to 2020. The impact of the financial incentives increase monotonically through the cohorts. Studying the impact of the financial incentives among partially disabled individuals who claimed benefits during the period from 2006 to 2013, [Koning and van Sonsbeek \(2017\)](#) found an average labor participation increase of 2.6 pp in the long-term. Our results show that this impact estimate is specific to the study period and has substantially changed over time.

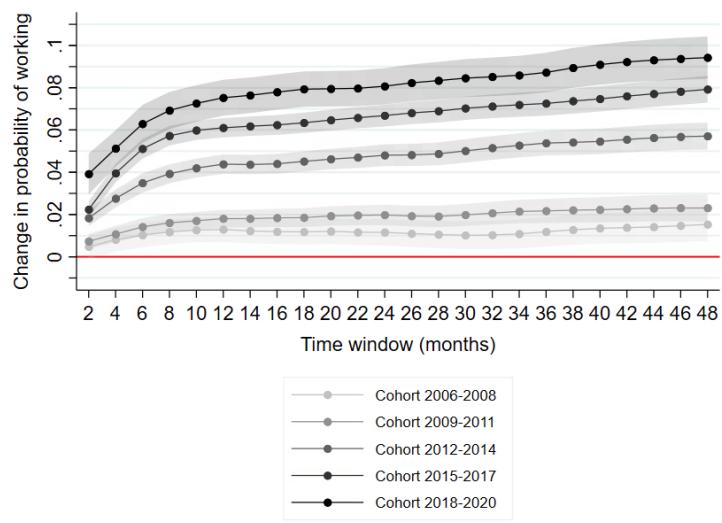


Figure 12: Estimated effects of financial incentives and 95% confidence intervals around them by year of entry to the DI scheme. Groups that enter the DI scheme in years later than 2006-2008 are re-weighted so that year of entry groups have similar distributions of age, gender, and contract type.

## Employment status

Employees who have a temporary contract and the unemployed may have different reintegration motivations than the employees who have a permanent contract. They may therefore respond to the financial incentives to different extents. As explained in Section 2, the financial incentives to resume work for partially disabled individuals are strong, but they are offered only after a period of up to 5 years and 2 months –claiming sickness benefits for 2 years and the first-stage wage-related disability benefit for up to 3 years and 2 months– complicating a successful return to the labor market. During the sickness period, individuals employed through temporary work agencies and the unemployed have no former employer to return to, and therefore they are likely to be unemployed when they face the financial incentives. In fact, in the study sample, 6 months prior to facing the financial incentives, while 41% of the permanent contract workers are observed to be not working, 74% of the temporary contract workers and 79% of the unemployed are observed to be not working. If without a job when subjected to the financial incentives, disabled individuals may struggle to resume working if job search costs are high, negotiating a suitable work schedule with an employer is difficult, adjusting non-work schedules is difficult, or if the offered wage in the new job is lower than the reservation wage (Koning and Lindeboom, 2015; Zaresani, 2020). These mean that financial incentives may be less effective among temporary contract workers and the unemployed. The much lower labor participation rates among the temporary contract workers and the unemployed prior to facing financial incentives may also imply substantial unused work capacity in these groups. Financial incentives may then increase labor participation more often in these groups.

Figure 13 presents the estimated effects of the financial incentives for permanent contract workers, temporary contract workers and the unemployed. Responses of permanent contract workers are much weaker than those of the other two groups. Moreover, across the wage groups of permanent contract workers, financial incentives do not show a wage gradient: Higher pre-sickness wages do not lead to stronger responses. These findings suggest substantial unused work capacity among temporary contract workers and the unemployed, and financial incentives are effective to induce these groups to utilize their remaining work capacity.

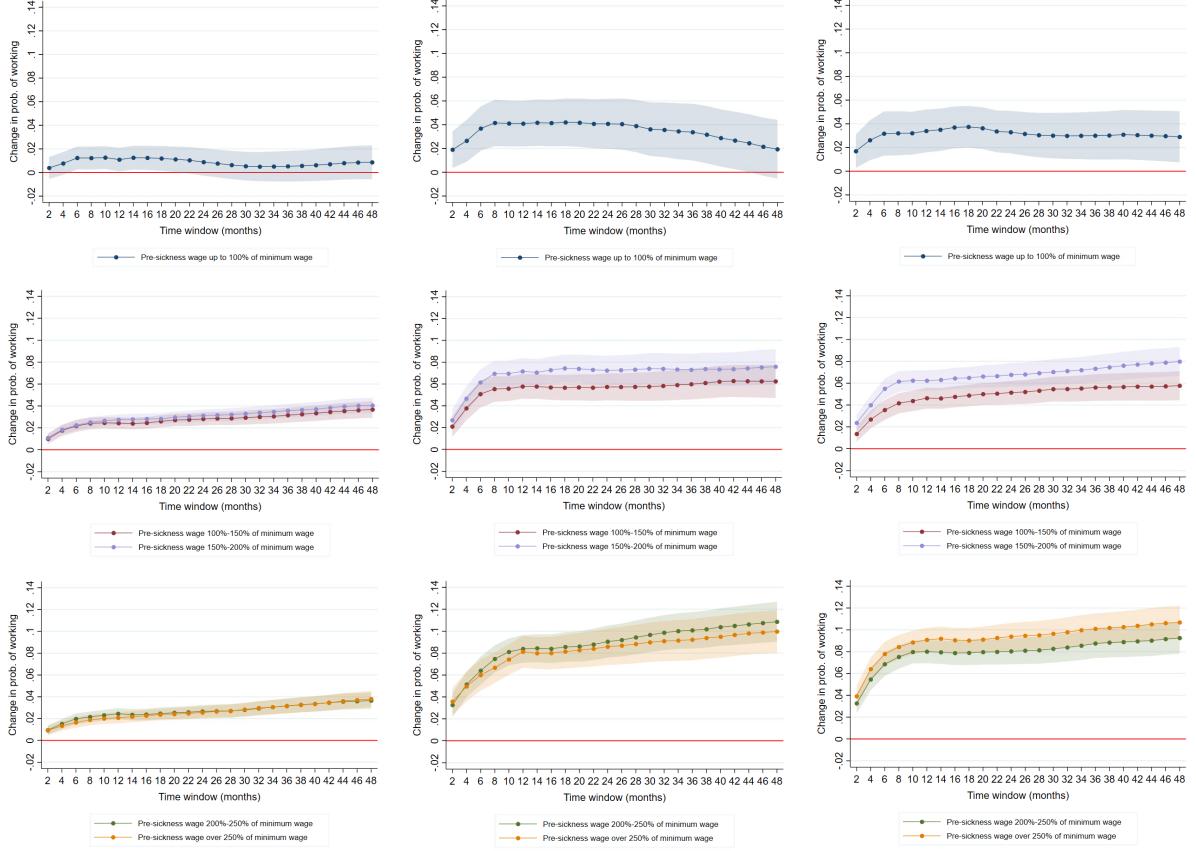


Figure 13: Estimated effects of financial incentives and 95% confidence intervals around them for individuals who have a permanent job (left panel), temporary job (center panel), and unemployed (right panel). Groups of individuals who have temporary jobs and who are unemployed are re-weighted so that the three employment status groups have similar distributions of age, gender, and year of entry to the DI scheme.

## Vacancies in the sector

For disabled individuals who have limited employment opportunities and hence a higher risk of unemployment, responding to the financial incentives to work can be more difficult. We consider the sectoral vacancy rate (the number of open vacancies per one thousand jobs) as an indicator of employment opportunities.<sup>15</sup> We distinguish two groups: individuals who at the time of reporting sick worked in sectors with vacancy rates below (e.g., agriculture, manufacturing, transport, health care, education, public sector) or above the median vacancy rate (e.g., construction, trade, financial services, catering). Figure 14 presents the results. For low-wage earners (up to 150% of minimum wage), labor market opportunities are not important for how disabled individuals respond to the financial incentives to resume work. This provides reinforcing evidence for our main finding that financial incentives are not effective for low income groups who face weaker financial incentives. In contrast, for disabled individuals who earn higher wages, financial incentives are more effective if labor market opportunities are better.

---

<sup>15</sup>The sector where sick individuals are or were employed is available in the disability data, and we determine the vacancy rate in each sector prior to and in the year of reporting sick using data from Statistics Netherlands.

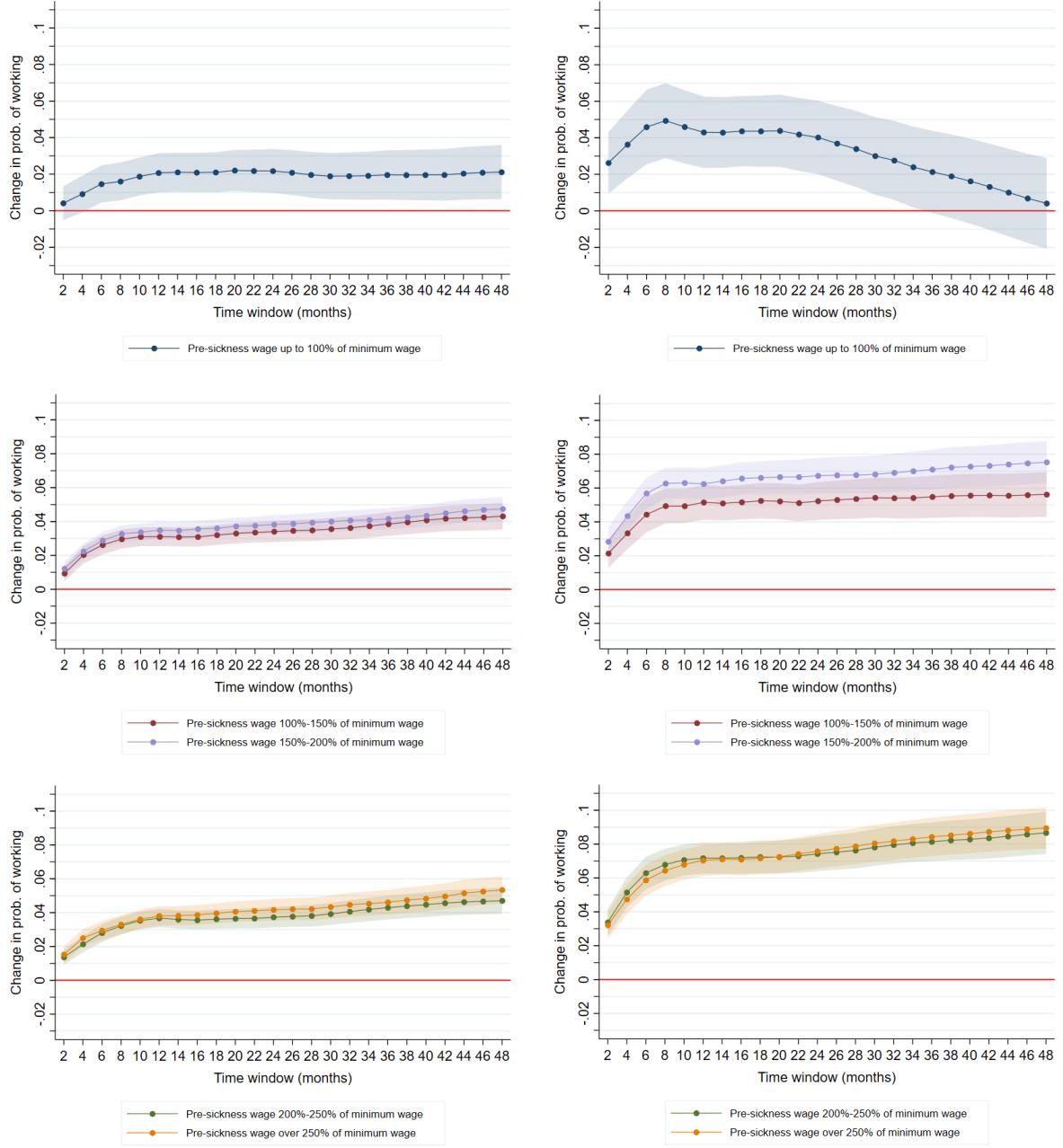


Figure 14: Estimated effects of financial incentives and 95% confidence intervals around them for individuals who work in a sector with vacancy rate above (top panel) and below (bottom panel) the median vacancy rate. The group who work in a sector with vacancy rate above the median is re-weighted so that the two groups of sectoral vacancy rate have similar distributions of age, gender, employment status, and year of entry to the DI scheme

## 7 Checking the identifying assumptions

### Window selection

The exclusion restriction assumption of the LR method requires that the potential outcomes do not depend on the value of the running variable inside a window, except via the treatment assignment indicator (Figure 6). The baseline results in Figure 7 show that the treatment effect estimate, as the estimate of the difference in the expected values of the potential outcomes, show an increase in the vicinity of the month the financial incentives take effect (time windows of 2 to 8 months). As discussed in Section 5.1, this can be due to, for example, individuals who are not able to react to the incentives immediately after the month the incentives take effect due to adaptation. Besides this increase, however, the treatment effect estimate is fairly stable through increasing time windows, suggesting that the potential outcomes do not depend on the value of the running variable, except through treatment assignment. Here we check to what extent the treatment effect estimate remains stable across wider time windows in a statistical sense.

We consider the p-value from a two-sample z-test of the difference between the treatment effect estimated considering a certain time window and that considering a benchmark time window of 10 months. The benchmark is determined based on the observation that, in Figure 7, treatment effect estimates start to stabilize when the time window is 10 months or wider for all wage groups. Figure 15 presents the p-value for all different time windows except for the time window of 10 months (benchmark for comparison). For time windows of sizes 8 to 22 months, the p-values for all wage groups are far larger than the conventional p-values for statistical significance. These results suggest that, in a statistical sense, the treatment effect estimate is stable across a wide range of windows around the cutoff providing evidence for the identifying exclusion restriction assumption.



Figure 15: P-values of the test of the difference in treatment effect estimates from different time windows and a time window of 10 months. Red reference line marks the p-value at 5%.

## Placebo cutoffs

The baseline results have shown that there are significant differences in labor participation around the cutoff, comparing individuals before and after they face the financial incentives. In order to assess that these differences are caused by the treatment and not by other factors, we compare the mean outcome around placebo cutoffs, before and after the true cutoff, keeping otherwise the regression specification the same as in the baseline. At placebo cutoffs before the true cutoff, this leads to a comparison that is between individuals who have not yet been treated. After the true cutoff, the comparisons are between treated individuals. Significant differences around placebo cutoffs may cast doubt on that the differences around the true cutoff are due to the treatment. In the context of the current study, however, it should be noted that the results at placebo cutoffs close to the true cutoff may be indicative of anticipation or adaption effects in response to the financial incentives.

Table 3 presents results based on placebo cutoffs at 30, 20 and 10 months before and after the true cutoff, using a time window of size 2 months. The are significant effects especially during the period before the true cutoff. The magnitudes of the placebo effects are, however, substantially smaller than the baseline effects at the true cutoff in Table 2. Furthermore, these significant effects seem to be closely related to the work resumption trends observed in Figure 9. The placebo effects visible at months before the true cutoff but not at months after it are potentially due to that work resumption is more likely early on during DI participation than later. Therefore, these results do not cast serious doubt on that the significant effects around the true cutoff are due to the financial incentives.

Table 3: Labor participation around placebo cutoffs

Pre-sickness wage as a fraction of the minimum wage	Placebo cutoff before the true cutoff			Placebo cutoff after the true cutoff		
	30	20	10	10	20	30
	months	months	months	months	months	months
<100%	0.008*	0.005*	0.004*	0.000	0.001	0.000
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
100–150%	0.005**	0.004**	0.004***	0.006***	0.002	-0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
150–200%	0.007***	0.003*	0.006***	0.003**	0.000	0.000
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
200–250%	0.005**	0.010***	0.003**	0.002*	0.002	0.000
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
>250%	0.006***	0.005***	0.005***	0.001	0.000	0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)

Notes: The window size is 2 months in each regression. The number of months before or after is in relation to the date of the true cutoff. \*\*\*, \*\*, \* denote statistical significance at 1%, 5% and 10%, respectively. Standard errors (in parentheses) account for heteroskedasticity and clustering at the individual level.

## Covariate balance

If local randomization assumptions hold, all factors driving the outcome variable other than the treatment indicator should not discontinuously change at the cut-off point. Although this assumption cannot be tested directly, relevant covariates can be checked for whether they change significantly at the cutoff. We consider health as a most relevant covariate. Information on medicines prescribed in a year, for which the costs are reimbursed under mandatory health care insurance (ZVV), are available from Statistics Netherlands. Based on this information, we calculate the number of medicines prescribed in a year, and consider larger number of medicines prescribed as indicator of worse health. We then estimate Equation (1) using this variable as the outcome. A significant treatment indicator suggests that treatment assignment is not random due to the outcome. Since the information on medicines prescribed is available on a yearly basis, it is difficult to test the impact of the treatment indicator based on (event) months, in particular at the cutoff month. That is, we need that values of the outcome can change at the cutoff month so that we can test whether the outcome exhibits a discontinuous jump at the cutoff month. To address this, we restrict the sample to individuals for who medicines are prescribed in different years before and at the cutoff month or later. For the smallest time window around the cutoff, for example, this implies considering individuals who are prescribed medicines in December of a year in the last month before the cutoff (event month -1) and in January of next year at the cutoff month (event month 0).

Figure 16 presents the estimated effects. Wider confidence intervals of estimated effects of wider time windows are due to smaller number of individuals having observations at those time windows. Number of medicines prescribed does not exhibit discontinuity at the cutoff in any wage group or time window except in rare cases. Note that, in our setting, it should be unlikely that covariates are imbalanced due to treatment since individuals face the treatment at random times depending on the number of years they worked before falling sick. Hence, any observed imbalance is likely be due to factors other than treatment.

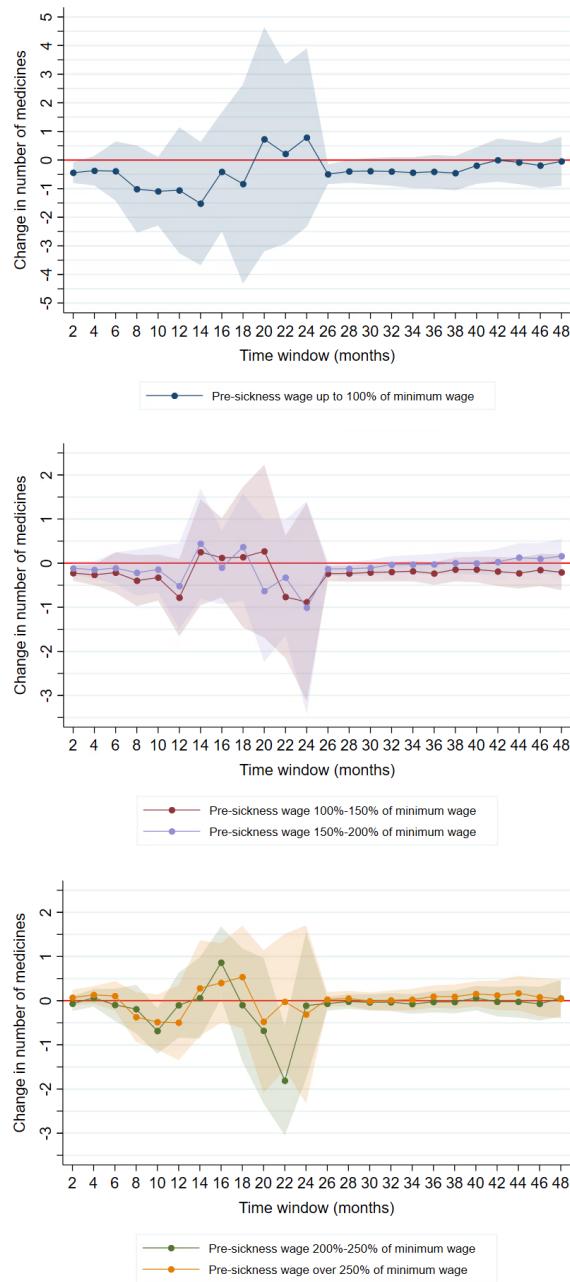


Figure 16: Estimated effects of financial incentives on number of medicines prescribed and 90% confidence intervals around them.

## Continuity-based approach to RD

We consider the widely implemented continuity-based approach to RD. Following standard practice, we consider a linear fit of the running variable, and a triangular kernel for weighting the observations centered around the cutoff. We select the window which minimizes the mean squared error (MSE) of the RD estimator. In addition, for inference, we perform the same estimation in a window which minimizes the coverage error (CER) of the confidence interval. Controls include dummies controlling for multiples of six months of DI receipt duration, and a linear function of the business cycle indicator. Table 4 presents the estimation results. The results provide evidence of an increasing effect of the financial incentives across groups earning higher wages and they are consistent with the results based on the LR approach to RD in Table 2. Table 5 presents the estimation results for social assistance receipt. Again, these results are consistent with the results based on the LR approach to RD in Figure 8.

As in Section 5.2, assuming that anticipation and adaptation are part of the true responses, we investigate to which extent they lead to an underestimation of the true effect of the financial incentives in the baseline regression analysis based on the continuity-based RD. Table 6 presents results based on the donut hole regression where we exclude observations within the time window of 8 months around the cutoff month, keeping otherwise the baseline regression specification the same. The treatment effect estimates are substantially larger than those based on the baseline regression. This suggests that baseline treatment effects might be underestimated due to anticipation and adaptation effects.

We repeat the checks on identifying assumptions conducted for the LR approach to RD in Section 7. Table 7 presents the estimated effects of financial incentives at placebo cutoffs, before and after the true cutoff, keeping otherwise the continuity-based RD specification the same as in the baseline. None of the placebo effects are significant, providing supporting evidence that the significant effects observed at the true cutoff are due to the financial incentives and not other factors.

If the continuity assumptions of the continuity-based approach to RD hold, the treatment should affect only the outcome at the cutoff and not covariates. To test this assumption, relevant covariates can be checked for whether they change significantly at the cutoff. We consider health as a most relevant covariate, and repeat the continuity-based RD estimation replacing the outcome with the number of medicines prescribed in the year concerned. A significant treatment indicator suggests that treatment assignment is not random due to the outcome. Table 8 shows that the treatment has no significant effect on the health outcome.

Table 4: Estimated effects of financial incentives on labor participation among pre-sickness wage groups based on continuity-based RD

Pre-sickness wage as a fraction of the minimum wage	Labor participation
<100%	0.019 (0.014)
100–150%	0.020*** (0.007)
150–200%	0.028*** (0.006)
200–250%	0.026*** (0.007)
>250%	0.033*** (0.007)

Notes: Regressions include the same covariates as in the baseline regression based on the LR approach. Standard errors are in parentheses. For each wage group, separate optimal bandwidths are estimated at the left and right of the cutoff. Bandwidths are chosen to minimize the mean squared error of the point estimate. Similar bandwidths are obtained when bandwidths are chosen to minimize the coverage error of the confidence interval. For labor participation, the bandwidths are 16.242, 11.451, 15.100, 11.781 and 13.092 at the left of the cutoff across wage groups in order of increasing wages. They are 35.446, 14.086, 12.350, 11.223 and 12.376 at the right of the cutoff.

Table 5: Estimated effects of financial incentives on social assistance receipt among pre-sickness wage groups based on continuity-based RD

Pre-sickness wage as a fraction of the minimum wage	Social assistance receipt
<100%	0.002 (0.006)
100–150%	0.004*** (0.001)
150–200%	0.004*** (0.001)
200–250%	0.005*** (0.001)
>250%	0.002*** (0.000)

Notes: Regressions include the same covariates as in the baseline regression based on the LR approach. Standard errors are in parentheses. For each wage group, separate optimal bandwidths are estimated at the left and right of the cutoff. Bandwidths are chosen to minimize the mean squared error of the point estimate. Similar bandwidths are obtained when bandwidths are chosen to minimize the coverage error of the confidence interval. The bandwidths are 20.321, 21.194, 22.500, 17.999 and 11.611 at the left of the cutoff across wage groups in order of increasing wages. They are 35.047, 28.762, 20.279, 32.423 and 28.546 at the right of the cutoff.

Table 6: Estimated effects of financial incentives on labor participation among pre-sickness wage groups based on donut hole continuity-based RD

Pre-sickness wage as a fraction of the minimum wage	Labor participation
<100%	0.037** (0.016)
100–150%	0.061*** (0.008)
150–200%	0.064*** (0.008)
200–250%	0.057*** (0.008)
>250%	0.074*** (0.008)

Notes: Regressions include the same covariates as in the baseline regression based on the LR approach. Standard errors are in parentheses. For each wage group, separate optimal bandwidths are estimated at the left and right of the cutoff. Bandwidths are chosen to minimize the mean squared error of the point estimate. Similar bandwidths are obtained when bandwidths are chosen to minimize the coverage error of the confidence interval. For labor participation, the bandwidths are 16.913, 12.902, 14.449, 12.179 and 13.572 at the left of the cutoff across wage groups in order of increasing wages. They are 26.593, 15.601, 20.946, 12.911 and 27.531 at the right of the cutoff.

Table 7: Labor participation around placebo cutoffs

Pre-sickness wage as a fraction of the minimum wage	Placebo cutoff before the true cutoff			Placebo cutoff after the true cutoff		
	30	20	10	10	20	30
	months	months	months	months	months	months
<100%	0.024 (0.031)	0.024 (0.031)	-0.002 (0.016)	0.002 (0.016)	0.007 (0.017)	-0.008 (0.019)
100–150%	0.004 (0.015)	0.004 (0.015)	0.001 (0.008)	0.003 (0.008)	-0.005 (0.008)	-0.002 (0.009)
150–200%	0.006 (0.012)	0.006 (0.012)	0.003 (0.007)	-0.004 (0.007)	-0.005 (0.008)	0.004 (0.008)
200–250%	0.003 (0.012)	0.003 (0.012)	0.000 (0.007)	-0.002 (0.008)	-0.002 (0.008)	-0.004 (0.009)
>250%	0.004 (0.011)	0.004 (0.011)	0.002 (0.007)	-0.005 (0.008)	-0.003 (0.008)	0.002 (0.009)

Notes: The window size is 2 months in each regression. The number of months before or after is in relation to the date of the true cutoff. \*\*\*, \*\*, \*, denote statistical significance at 1%, 5% and 10%, respectively. Standard errors (in parentheses) account for heteroskedasticity and clustering at the individual level.

Table 8: Estimated effects of financial incentives on the number of medicines prescribed among pre-sickness wage groups based on continuity-based RD

Pre-sickness wage as a fraction of the minimum wage	Number of medicines prescribed
<100%	-0.412 (0.375)
100–150%	-0.202 (0.179)
150–200%	-0.108 (0.163)
200–250%	0.041 (0.179)
>250%	0.013 (0.179)

Notes: Regressions include the same covariates as in the baseline regression based on the LR approach. Standard errors are in parentheses. For each wage group, separate optimal bandwidths are estimated at the left and right of the cutoff. Bandwidths are chosen to minimize the mean squared error of the point estimate. Similar bandwidths are obtained when bandwidths are chosen to minimize the coverage error of the confidence interval. For labor participation, the bandwidths are 19.383, 17.496, 20.359, 25.737 and 18.731 at the left of the cutoff across wage groups in order of increasing wages. They are 30.600, 37.772, 40.713, 35.916 and 39.420 at the right of the cutoff.

## An alternative identification strategy

To analyze the causal impact of the financial incentives of the DI scheme for partially disabled individuals, Koning and van Sonsbeek (2017) use a different identification strategy than the one we use. Exploiting the variation in the time individuals spend receiving the wage-related benefit, they compare the labor participation of “treated” individuals who are close to or after they face the financial incentives with those of the “control” group of individuals who are not or not yet affected by the financial incentives. This identification strategy relies on two assumptions. First, treatment and control group individuals share the same time trend in the potential outcome before and after they face the financial incentives or not. Second, relevant behavioral changes occur in a specific time window around the month financial incentives take effect. Their regression model includes two dummies that capture the short- and long-term effects of the financial incentives. One dummy indicates a time window (6 months in the benchmark model) around the month financial incentives take effect, and another dummy indicates the succeeding months. Other preceding months are chosen as base. Coefficients on these dummies represent the treatment effects. Other controls include five-year age group dummies interacted with a fourth order polynomial of the time spent claiming DI, calendar year dummies, and individual fixed effects.

We estimate the regression specification outlined above using our data, and compare the resulting treatment effect estimates with those based on the regression specified in Equation (1). We then check if both identification strategies lead to the same qualitative conclusions. Table 9 presents the estimated short- and long-term effects of the financial incentives. These estimates are close to the estimates from small and large window sizes in Figure 7. This suggests that an alternative identification strategy that relies on different identifying assumptions confirm the results based on the RD method we have used.

Table 9: Results based on the identification strategy of Koning and Van Sonsbeek

Pre-sickness wage as a fraction of the minimum wage	Labor participation	
	Short-term	Long-term
<100%	0.008 (0.007)	0.009 (0.010)
100–150%	0.029*** (0.004)	0.044*** (0.005)
150–200%	0.028*** (0.003)	0.052*** (0.005)
200–250%	0.032*** (0.003)	0.061*** (0.005)
>250%	0.025*** (0.003)	0.062*** (0.005)

Notes: Regressions include the same covariates as in the baseline regression based on the LR approach. Standard errors are in parentheses. Number of observations are 176,009, 786,929, 906,612, 793,063 and 719,541 used in each regression are for the five wage groups, respectively across wage groups in order of increasing wages. Number of individuals are 2,496, 10,454, 12,318, 10,839, and 10,143.

## **Duration of DI receipt**

In the baseline specification, to account for the possible effect of the duration of DI receipt, we considered dummies controlling for multiples of six months of DI receipt duration. A stricter measure of the elapsed duration could affect the treatment effect estimate if treatment is closely correlated with the running variable within a time window around the cutoff. Figure 17 presents results when we allow for dummies controlling for multiples of three months of DI receipt duration in the regression. Compared to the baseline results in Figure 7, the main difference is that the magnitudes of the estimated treatment effects are somewhat smaller. We conclude that our baseline results are fairly robust to how we control for duration of DI receipt.

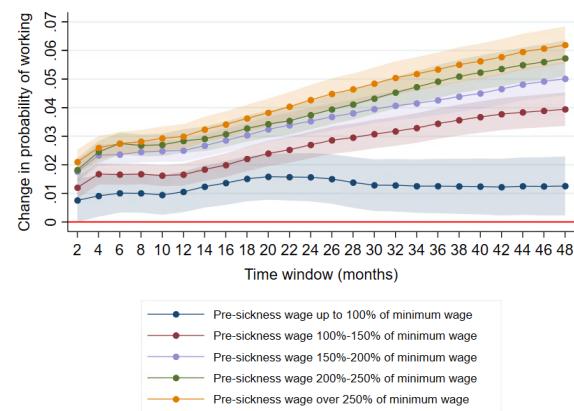


Figure 17: Estimated effects of financial incentives and 90% confidence intervals around them based on regressions with dummies that control for multiples of three months of DI receipt duration.

## 8 Conclusion

Policy implications and recommendations of our study are four-fold. First, the main gains of the DI reforms in the past two decades have been reached by stricter screening during the sickness period and tighter eligibility criteria during the DI period. Our results suggest that financial incentives built into the DI scheme add to this significantly.

Second, the financial incentives are more effective for individuals at the middle to higher end of the earnings distribution. This implies that the financial incentives cannot be relaxed for these individuals without causing adverse effects in terms of lower labor participation and earnings. On the other hand, financial incentives are much less effective for more than a quarter of DI recipients who are low-wage earners. Their incentive to resume working is comparatively low, and, on top of their health impairment, lack of skills and job opportunities may put additional constraints on their labor participation. Since benefits are already near the social minimum level for a large part of this group, the possibilities for increasing the financial incentives for work seem limited.

Third, the disparity in how income groups respond to the financial incentives suggests that the incentives might unintentionally increase inequality. Although the higher wage groups have more to lose from the stronger financial incentives, they are also better able to compensate a disability related income loss through alternative income sources.

Finally, a policy recommendation regards how DI beneficiaries are informed of the financial incentives of the DI scheme. A beneficiary receives two letters which inform about the financial incentives; one at the moment the benefit is awarded and another shortly before expiration of the first-stage wage-related benefit. However, neither of the two letters explicitly informs about the size of the incentive and the difference between the two types of second-stage benefits, nor they explain how working more, below and above the 50% threshold, affects the benefit level and the total income from wage and benefit. By making this information explicit, beneficiaries can better anticipate the incentives to come so that effectiveness of the incentives can be improved.

## References

- Arulampalam, W., 2001. Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal* 111 (475), F585–606.
- Arulampalam, W., Gregg, P., Gregory, M., 2001. Introduction: unemployment scarring. *The Economic Journal* 111 (475), F577–584.
- Autor, D. H., Duggan, M., Greenberg, K., Lyle, D. S., 2016. The impact of disability benefits on labor supply: Evidence from the VA's disability compensation program. *American Economic Journal: Applied Economics* 8 (3), 31–68.
- Borghans, L., Gielen, A. C., Luttmer, E. F. P., 2014. Social support substitution and the earnings rebound: evidence from a regression discontinuity in disability insurance reform. *American Economic Journal: Economic Policy* 6 (4), 34–70.
- Bound, J., Burkhauser, R. V., 1999. Chapter 51 economic analysis of transfer programs targeted on people with disabilities. Vol. 3 of *Handbook of Labor Economics*. Elsevier, pp. 3417–3528.
- Burkhauser, R., Marc, D., McVicar, D., Wilkins, R., 2014. Disability benefit growth and disability reform in the US: lessons from other OECD nations. *IZA Journal of Labor Policy* 3 (4), 1–30.
- Butler, B., Volkerink, M., Lung, C. L., 2019. Nieuwe conjunctuur indicator voorspelt afvlakking groei in 2019. *Economisch Statistische Berichten* 104 (4773).
- Bütler, M., Deuchert, E., Lechner, M., Staubli, S., Thiemann, P., 2015. Financial work incentives

- for disability benefit recipients: lessons from a randomised field experiment. *IZA Journal of Labor Policy* 4 (1), 1–18.
- Campolieti, M., 2004. Disability insurance benefits and labor supply: some additional evidence. *Journal of Labor Economics* 22 (4), 863–889.
- Campolieti, M., Riddell, C., 2012. Disability policy and the labor market: evidence from a natural experiment in Canada, 1998–2006. *Journal of Public Economics* 96 (3–4), 306–316.
- Cattaneo, M. D., Frandsen, B. R., Titiunik, R., 2015. Randomization inference in the regression discontinuity design: An application to party advantages in the U.S. senate. *Journal of Causal Inference* 3 (1), 1–24.
- Cattaneo, M. D., Idrobo, N., Titiunik, R., 2020. *A Practical Introduction to Regression Discontinuity Designs: Foundations. Elements in Quantitative and Computational Methods for the Social Sciences*. Cambridge University Press.
- Cattaneo, M. D., Titiunik, R., 2021. Regression discontinuity designs. Mimeo.
- Favre, G., Haller, A., Staubli, S., 2021. Induced entry in disability insurance: evidence from Canada. National Bureau of Economic Research, Center paper NB20-08.
- Gruber, J., 2000. Disability insurance benefits and labor supply. *Journal of Political Economy* 108 (6), 1162–1183.
- Hainmueller, J., 2012. Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20 (1), 25–46.
- Halla, M., Schmieder, J., Weber, A., 2020. Job displacement, family dynamics, and spousal labor supply. *American Economic Journal: Applied Economics* 12 (4), 253–287.
- Haller, A., Staubli, S., Zweimüller, J., 2020. Designing disability insurance reforms: Tightening eligibility rules or reducing benefits. National Bureau of Economic Research Working Paper Series 27602.
- Kantarcı, T., van Sonsbeek, J.-M., Zhang, Y., 2023. The heterogenous impact of stricter criteria for disability insurance. *Health Economics*, 1–23.
- Karlström, A., Palme, M., Svensson, I., 2008. The employment effect of stricter rules for eligibility for di: Evidence from a natural experiment in sweden. *Journal of Public Economics* 92 (10–11), 2071–2082.
- Koning, P., Lindeboom, M., 2015. The rise and fall of disability insurance enrollment in the Netherlands. *Journal of Economic Perspectives* 29 (2), 151–172.
- Koning, P., van Sonsbeek, J.-M., 2017. Making disability work? The effects of financial incentives on partially disabled workers. *Labour Economics* 47, 202–215.
- Kostøl, A. R., Mogstad, M., 2014. How financial incentives induce disability insurance recipients to return to work. *American Economic Review* 104 (2), 624–655.
- Malani, A., Reif, J., 2015. Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform. *Journal of Public Economics* 124, 1–17.
- Mattei, A., Mealli, F., 2017. Regression discontinuity designs as local randomized experiments. *Observational Studies* 3 (2), 156–173.
- Mullen, K. J., Staubli, S., 2016. Disability benefit generosity and labor force withdrawal. *Journal of Public Economics* 143, 49–63.
- OECD, 2023. Public spending on incapacity (indicator), Accessed on 29 May 2023.
- Ruh, P., Staubli, S., 2019. Financial incentives and earnings of disability insurance recipients: evidence from a notch design. *American Economic Journal: Economic Policy* 11 (2), 269–300.
- Sales, A. C., Hansen, B. B., 2020. Limitless regression discontinuity. *Journal of Educational and Behavioral Statistics* 45 (2), 143–174.
- Staubli, S., 2011. The impact of stricter criteria for disability insurance on labor force participation. *Journal of Public Economics* 95 (9–10), 1223–1235.

- Vall Castelló, J., 2017. What happens to the employment of disabled individuals when all financial disincentives to work are abolished? *Health Economics* 26 (S2), 158–174.
- Van Sonsbeek, J.-M., Gradus, R. H. J. M., 2013. Estimating the effects of recent disability reforms in the Netherlands. *Oxford Economic Papers* 65 (4), 832–855.
- Weathers, R. R., Hemmeter, J., 2011. The impact of changing financial work incentives on the earnings of social security disability insurance (ssdi) beneficiaries. *Journal of Policy Analysis and Management* 30 (4), 708–728.
- Zaresani, A., 2020. Adjustment cost and incentives to work: Evidence from a disability insurance program. *Journal of Public Economics* 188 (104223).

## Appendix

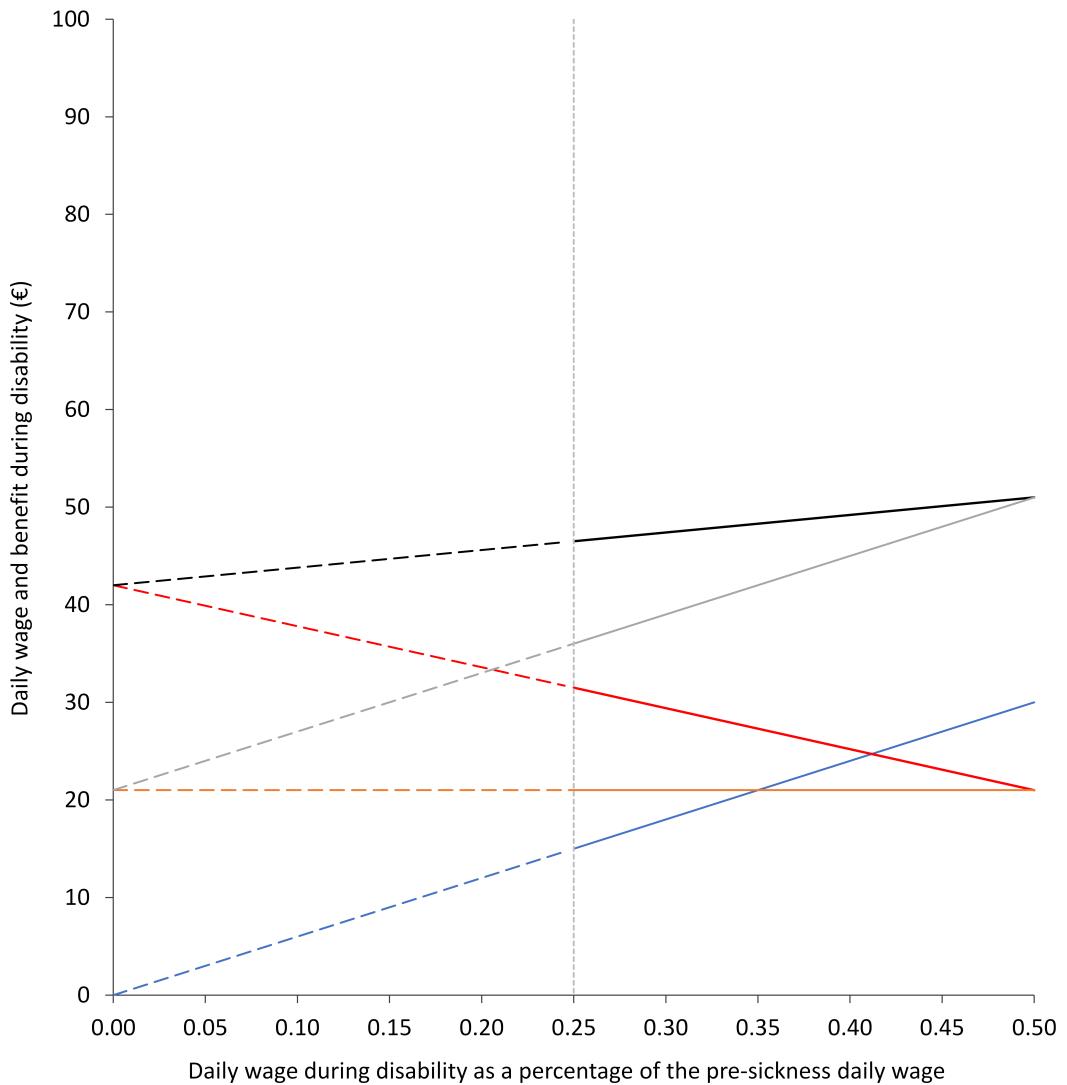


Figure 18: Financial incentives to work in the second stage of the DI scheme for an individual who earns €60 before falling sick and has a disability grade of 50%. Blue: Wage. Red: First-stage (wage-related) benefit. Orange: A second-stage (follow-up or wage supplement) benefit. Black: Sum of wage and the first-stage benefit. Gray: Sum of wage and a second-stage benefit. Vertical reference line: Remaining work capacity utilization rate = 50%. Dashed lines apply when remaining work capacity utilization rate < 50%. Solid lines apply when remaining work capacity utilization rate  $\geq 50\%$ . The financial incentive is the difference between the dashed black line (sum of wage and the wage-related benefit) and dashed gray line (sum of wage and the follow-up benefit) if remaining work capacity utilization rate < 50% in the second stage of the DI scheme (left side of vertical reference line). The financial incentive is the difference between the solid black line (sum of wage and the wage-related benefit) and the solid gray line (sum of wage and the wage-supplement benefit) if remaining work capacity utilization rate  $\geq 50\%$  in the second stage of the DI scheme (right side of vertical reference line).

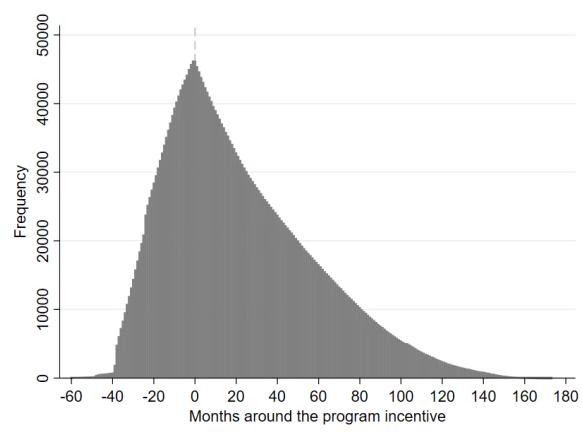


Figure 19: Number of months spent in the DI scheme before (first stage of the DI scheme) and after (second stage of the DI scheme) the month financial incentives take effect.

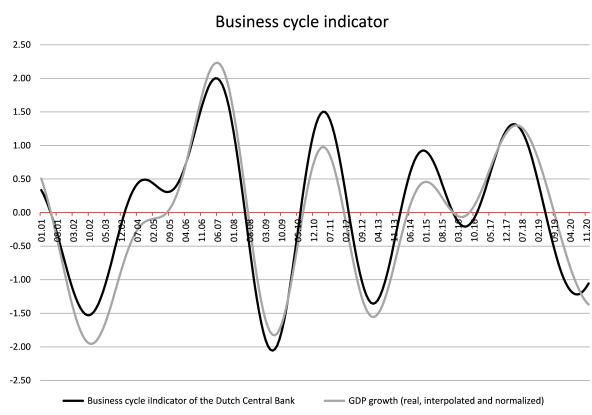


Figure 20: Business cycle indicator of the Dutch Central Bank.

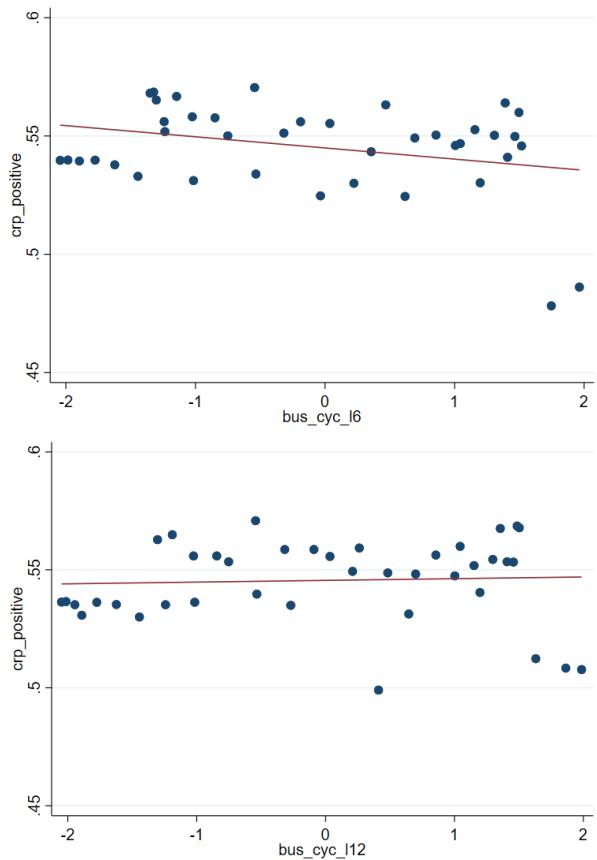


Figure 21: Binned scatter plot of mean labor participation against the business cycle indicator. Each dot represents the mean labour participation for that bin of the business cycle indicator. The fitted line is based on regression using the underlying data that make up the bins. The outcome is labor participation and the only covariate is the business cycle indicator. The top and bottom panels consider lags of 6 and 12 months of the business cycle indicator, respectively.